# THE REAL EFFECTS OF OPERATIONAL RISK: EVIDENCE FROM DATA BREACHES

Matteo Binfarè

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Finance at the Kenan-Flagler Business School.

Chapel Hill
2020

Approved by:

Gregory W. Brown

Robert Connolly

Robert S. Harris

Christian T. Lundblad

Paige Ouimet

**ABSTRACT**

Matteo Binfarè: The Real Effects of Operational Risk: Evidence from Data Breaches
(Under the direction of Gregory W. Brown)

Operational risk poses unique challenges to corporations around the world. However, little is known about the consequences of operational risk on financing costs and firm outcomes. In this paper, I document substantial and persistent effects on financing costs and debt contracting caused by operational risk following data breaches of public firms. Exploiting data breach events between 2005 and 2015, I find that lenders charge breached firms 15 to 20 percent larger spreads, and tighten covenant intensity, consistent with a shift in control rights over cash flows. The effect is larger for breaches of financial information or malicious cyber-attacks, and for firms with lower attention to risk management. Moreover, financial and operating leverage increases, profitability drops, and firms face a higher probability of default. Lastly, *ex*-ante mispricing by banks does not explain these findings.

To my Italian family, to my American family, and to the Pocket family for the continuous support and love

**ACKNOWLEDGEMENTS**

I am grateful to my committee for continuous support and mentoring: Gregory W. Brown (Chair), Robert S. Harris, and Paige Ouimet. I am also indebted to Yasser Boualam, Ric Colacito, Robert Connolly, Jennifer Conrad, Jesse Davis, Andra Ghent, Andrei Gonçalves, Yunzhi Hu, Camelia Kuhnen, Christian Lundblad, Adam Reed, Jacob Sagi, and Elena Simintzi for helpful guidance and valuable insights. I thank Isil Erel, Steve Karolyi (discussant), Greg Nini, and seminar participants at the 2019 Northern Finance Association Meetings & UNC for helpful comments and suggestions. All errors are my own.

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

CHAPTER 1

## 1.1 Introduction

While the asset pricing literature has identified several sources of risk, corporate finance research struggles to understand which type of risk matters and the value of risk management. In the context of market risk, Campello et al. (2011) show that corporate hedging alleviates frictions in the bank loan market. While active market and credit risk management can mitigate losses stemming from negative outcomes (Brown and Toft (2002); Carter et al. (2006); Jin and Jorion (2006); Smith and Stulz (1985)), negative shocks to a firm's *intangible* capital can have adverse and unexpected effects that corporations can only partially hedge or anticipate. Operational risk (i.e. risk of losses from inadequate internal processes, people and systems, or external events) has received substantial attention by firms and regulators alike, yet little is known about its effects on borrowing costs and firm value.

In this paper I provide the *first* evidence of increased financing costs and changes in debt contractual features stemming from operational risk vulnerabilities at publicly listed firms. I exploit negative shocks to a firm's stock of digital intangible capital through data breaches due to limitations in operational risk management. I identify mechanisms through which operational risk management vulnerabilities affect a firm cost of capital, and ultimately firm value. Specifically, I document substantial and persistent effects on financing costs and debt contracting following data breaches of publicly listed firms.

First, private lenders (e.g. banks) charge breached firms 15 to 20 percent larger spreads, compared to similar non-breached firms. Banks also tighten covenant intensity and are more likely to include covenants, consistent with a shift in control rights over cash flows.

Second, I find interesting cross-sectional heterogeneity. Specifically, consistent with economic tensions between the value of data (i.e. digital assets) and investments in risk management, the effect is stronger when the breach involves customer financial information or comes through a malicious third-party entry (cyber). More visible firms (as proxied by Fortune 500 indicator), without a Chief

Information Officer (CIO), and larger information asymmetries before a breach, pay significantly larger spreads.

Third, data breaches also impact firm characteristics. Specifically, firm financial and operating leverage increases, profitability declines (ROA), and the likelihood of a second breach increases. Data breaches also impact default risk, since they reveal time-invariant risk management flaws that impact tail risk and negatively affect firm operations going forward. Finally, I show that mispricing of loans by lenders does not explain the post-breach increase in financing costs. Evidence suggests that banks update upon observing a breach to reflect firm riskiness going forward, rather than realize that borrowing firms received lenient pricing terms before a breach.

The private loan market represents an ideal laboratory for investigating the interplay between risk management and firm financing costs in the context of a new type of operational risk. It is important to understand the implications of data breaches for financing costs as they convey information about firm risk that has a first-order effect on debt contracts. First, corporate financing by banks via the loan market is the largest source of funding for U.S. firms. For instance, banks provided $2.8 trillion of credit to U.S. firms in 2018, while approximately $1 trillion in 1998. Chakraborty et al. (2018) estimate that bank debt accounts for nearly 50% of a firm debt structure. Second, because of the private nature of the syndication process, lead arrangers often have a clearer picture of the future cash flows, material risks, and management actions of a firm. In addition, creditors can quickly amend the contractual features of a loan to accommodate changes in fundamentals, and shift the balance of ownership and control in their favor.

Data breaches have made the headlines in recent years. According to a centralized global database, 1,765 breaches occurred in 2017, with more than 2.5 billion identities exposed globally. In the first half of 2018, 3.5 billion individual records were compromised, for a total of 944 incidents.[1] Despite massive worldwide spending on information security and risk management, sectors ranging from consumer discretionary, banking, and health care, to manufacturing have been targeted, with the theft of digital data from many high-profile public companies (such as Capital One, CVS, Facebook, Marriott, and Target). United States firms spend the most on post-data breach response, with

[1]https://breachlevelindex.com/data-breach-library

the overall U.S. (global) costs averaging $8.19 (3.92) million *per* breach.[2] Good operational risk management should in theory decrease the likelihood of negative outcomes, thus decreasing the probability of default. In a world of *active* risk management, if risks can be perfectly hedged and predicted with certainty, firm value should not change in response to extreme events, after accounting for consumer reactions and regulatory fines. However, if investors update the magnitudes and the probabilities of negative events, then firm value should change.

Staggered shocks to digital assets through data breaches provide a clean laboratory for understanding the impact of operational risk management vulnerabilities on borrowing costs and firm value. Because of their plausibly exogenous nature, data breaches represent a useful testing ground for understanding how operational risk ripples through firms by impacting default risk and firm outcomes. First, all firms are under constant attack and very few appear to be immune from breaches. However, conditional on being a likely target, the timing of a breach is *exogenous*. Second, data breaches convey little information about firms' *status quo* or their products, as investors can collect information about management competence and economic conditions through other sources. Third, a careful matching procedure combined with a difference-in-difference empirical strategy and the staggered nature of breaches over time, help alleviate concerns that a time-varying omitted variable confounds the results. Moreover, contextualizing data breaches within the borrower-lender relationship allows the study of the response of banks to an *exogenous* shock to their portfolios. This is important because firms do not set their interest rates, but rather banks make informed decisions after observing a breach and quantifying the marginal contribution that a new loan would add to the risk of their portfolios. From the bank's perspective, a data breach signals inherent risks which ripple through firm operations, and affect firm outcomes and default risk going forward in a way that requires contract features to change. Lastly, data breaches represent a unique setting compared to other types of operational risks an enterprise faces: They pose systematic and entity-specific risks at the same time, and the process governing the probability of a breach is persistent over time and across industries. Moreover, the value of data conflates with internal controls and investment in risk management in unique ways (compared to fraud, natural disaster, or misreporting).

[2] https://www.ibm.com/security/data-breachIBM 2019 Cost of a Data Breach Report

To study these effects on loan contracting and firm outcomes, I first collect comprehensive information on data breaches at publicly listed companies (or subsidiaries of public companies) over the period 2005-2015. The data comes from the Privacy Rights Clearinghouse's Chronology of Data Breaches (PRC), and contains details about the breached organization (private, public, nonprofit, government, etc.), the type of data breach, the number of records stolen, the date of discovery, and a description of the incident. PRC collects information from several sources, such as the media, newspapers, SEC filings, United States Attorney Generals, government agencies, and IT security websites. I first document the nature of data breaches, and describe the firms that suffer a data breach. Affected firms tend to be large, profitable and with more investment opportunities (as measured by Tobin's $q$).

Second, I provide novel evidence that private lenders (e.g. banks) respond by charging higher spreads to breached borrowers. Using a difference-in-difference framework, I find that breached firms pay 15 to 20 percent higher spreads than similar non-breached firms. These results are robust to a variety of empirical specifications and tests, including two placebo tests and the inclusion of additional controls and fixed effects. I document economically important effects: Given an average duration of debt of breached firms of about 4 years, I estimate that the interest rate increase would translate into a 1.25 to 1.75 percent loss in debt value – provided the increase in financing costs is shared across maturities and debt types. I also show that the effect of increased spreads is more pronounced among firms that suffer a breach of financial information, such as loss of credit and debit cards numbers, online platform credentials, and bank account information (financial), as well as among data breaches through malicious third-party entry (cyber). Similarly, repeated breaches, as proxied by the total number of data breaches suffered by a firm over the sample period, and breaches with more records compromised, require larger spreads.

Third, I show that in line with the asymmetric payoff structure of debt holders, banks promptly alter the covenant composition of debt contracts and set their pricing terms accordingly. If firm default risk increases, then preventive covenants play an important role. In the context of data breaches, banks anticipate that either shareholders will undergo costly investments (likely via financial or operating leverage) or will make poor use of excess cash flows. My empirical findings suggest that creditors rely on general covenants more extensively than before. In addition, firms do not pledge assets as collateral more often, but covenant tightness and intensity (i.e. the total number of general covenants included in a contract) increase. The effect is substantial and translates to a 15 percent higher likelihood of loans with general covenants, and a 20 percent increase in covenant tightness. These effects have real-side implications by shifting part of the control rights over free-cash flows to debt holders and restricting firm actions.

Fourth, I look at the heterogeneous cross-sectional responses by breached firms and find that the effect of higher interest rates is concentrated among large, highly visible and reputable, yet vulnerable firms, as measured by the presence of a cyberrisk committee on the board. This is in line with the idea that data breaches are in fact a negative shock to otherwise healthy firms. Banks update their assessment of operational risk management and distress risk going forward.

I then explore the likely channels through which the changes in debt contractual terms may arise. I show that both the probability of distress (and measured by the distance to default estimated via Merton's structural model) and the conditional probability of a data breach increase. These two facts are consistent with the idea that data breaches alter the probability distribution of distress, giving debt holders more power during the syndication process. In addition, return on assets decreases (although with no statistical significance at conventional levels), likely indicating loss of customers, extraordinary items, or increased interest expenses (or a combination). Lastly, firms raise their *financial* and *operating* leverage via an increase in total debt and operating leases. Operating leases are risky for debt holders, as they represent fixed recurring payments. However, most short-term operating leases represent investments in IT systems, data centers, or cloud computing. Consistent with material changes in the capital structure of borrowing firms, banks charge higher interest rates on loan arrangements with greater probability of capital covenant violation at inception.

Finally, data breaches may indicate that loans were incorrectly priced and banks update upon observing a breach. In this case, the increase in interest rates comes from banks repricing loans to higher levels consistent with each borrower's adequate risk-profile. I conduct a series of tests to refute this alternative hypothesis, and corroborate the view based on changes in economic fundamentals following an attack. This is also broadly consistent with the existing literature (see for example Akey et al. (2018); Kamiya et al. (2018)).

My paper contributes to several strands of the literature. First, I show that operational risk management is valuable using a quasi-exogenous shock to the intangible assets that risk management should protect. Second, while the previous literature on operational risk failures (e.g. fraud, data breaches, lawsuits, natural disasters, etc.) suggests that shareholder and investor wealth decreases following such events, I provide evidence of value losses stemming from the financing side of the economy. Existing computer science and operations literature suggests that firms suffer abnormal stock market drops of about 0.5 to 2 percent at the time of announcement (Campbell et al. (2003); Cavusoglu et al. (2004); Gatzlaff and McCullough (2010); Gordon et al. (2011)). Spanos and Angelis (2016) analyze 37 articles related to the stock market reaction to security breaches, phishing, and other vulnerabilities. The authors find that 90% (70%) document a negative (significant) effect. Many of the existing studies argue that loss of reputation, potential fines, and negative effects on consumers likely explain the stock market response. I depart from this literature by focusing on the long-term effects of data breaches on debt holders rather than shareholders of target firms. The negative market reaction to unexpected data breaches is well documented, while lenders' response has been largely unexplored. In view of the asymmetric nature of creditors' payoff structure, it seems natural to investigate the implications of risk management for debt contracting. More recently, Kamiya et al. (2018) show that cyber-attacks are correlated with changes in firm's outcome, risk management practices, and CEO compensation. Similarly, Akey et al. (2018) report an increase in corporate social responsibility investments (CSR) for target firms after a data breach. They also document valuation losses consistent with damages to the reputation of a firm. I add to this literature by providing new evidence of the *real* and *financial* costs of a data breach. The private loan market allows me to analyze how lenders respond and reassess the risk-profile of borrowers in their loan portfolio, after negative shocks that convey increased default risk. Although there are other costs related to data breaches (e.g. regulatory fines), I do not explicitly consider them here. While it is true that regulatory fines can cost firms million of dollars, they have become more common only after the GDPR requires direct actions against breached firms.

Finally, I add to the literature on bank loans and corporate outcomes after negative shocks to firms. For example, Gormley and Matsa (2011) show that firms respond to increased liability risk by undertaking value-destroying corporate acquisitions. Similarly, Yuan and Zhang (2015) and Deng et al. (2014) study the effect of litigation risk and shareholder lawsuits on the pricing and non-pricing terms of bank loans. They show that banks charge higher spreads after such events and use tighter covenants, consistent with a loss in reputation. Graham et al. (2008) and Chava et al. (2017) study the effect of financial misreporting and restatement on bank loans. In particular, Chava et al. (2017)

shows that firms suffer a loss in reputation which is difficult to rebuild. I complement and add to this literature by considering a novel shock to *intangible* capital, possibly unrelated to managerial skills, firm's products or accounting quality, and by suggesting that the actual costs of data breaches on bank lending are as large as those documented by previous studies.

The paper proceeds as follow. Section 1.2 provides summary statistics, data sources, and the empirical strategy. Section 1.3 examines the pricing effect of data breaches, and the heterogeneous response by breach type. Section 1.4 focuses on the non-pricing terms, while Section 1.5 looks at the cross-sectional heterogeneity by borrower characteristics. Section 1.6 studies the likely mechanisms and alternative explanations, while Section 1.7 describes additional robustness tests. Section 1.8 concludes.

## 1.2 Data and Summary Statistics

### 1.2.1 Data Sources and Sample

I obtain data on breach events from the Privacy Rights Clearinghouse's Chronology of Data Breaches (henceforth PRC).[3] PRC maintains a chronology of data breaches between 2005 and the present and provides information on the type of breach (payment card fraud, hacking, or unintended disclosure, etc.), the type of organization, the number of affected records, and a brief description of the incident. Since 2005, 9,046 data breaches have been made public. These amount to 11,600,939,373 records breached.

The initial sample consists of 4,880 data breaches over the period 2005 to 2015 in both public and private entities (therefore excluding breaches to nonprofits, government agencies, educational institutions, etc.). Panel A of Table 1.1 shows that most data breaches occur through hacking (23%), unintended disclosure of digital information (16%), and physical or portable device with electronic information (42%). The bulk of the breached data includes general information about privacy such as names, emails, addresses and login credentials, while one-fifth of data breaches includes loss of financial information, such as debit and credit cards. The average number of records stolen was 640,940.

[3] See https://www.privacyrights.org/data-breaches

I identify data breaches that occurred at publicly listed companies or at subsidiaries of publicly listed companies using CRSP, Compustat, and ORBIS. This method leads to 386 events at 230 unique firms. A fourth of the firms experience only one type of data breach during the sample period, while the remaining 75 percent experienced more than one. Of the 230 unique firms, the average number of records breached was 2,052,698, while the average total number of breach events was 1.6 per firm.

Figure 1.1 plots the total number of data breaches in the three most common categories between 2005 and 2015. Figure 1.1 shows that the number of data breaches per year increased over time, reaching its peak at the onset of the financial crisis of 2008-2009 and in 2014, amid the equity bull market. Furthermore, the type of data breach appears to trend over time, and it is likely driven by technological advances, more digitally stored information, and state-level disclosure legislation.

Stock prices and firm characteristics are obtained from CRSP and Compustat, respectively. I require firms to have non-missing balance sheet and stock price data, as well as bank loans in the years before and after a data breach. This filtering procedure leaves me with 122 unique firms for a total of about 1,000 loans.

The bank loan data come from the WRDS-Thomson-Reuters' LPC DealScan. The DealScan database contains comprehensive historical information on pricing of loans and details, terms, and conditions for debt contracts. DealScan gathers data from SEC filings and other publicly available sources (10Ks, 10Qs, 8Ks, Exhibits 10), debt capital market syndicates and bankers, as well as other internal records. I obtain information at the facility and at the package level, consisting of the loan spread over LIBOR, loan size, loan maturity, covenants, number of lenders per syndicate, performance pricing provision and collateral use. I refer to the Chava and Roberts (2008) Dealscan-Compustat Linking Database provided by Professor Michael Roberts to match between loans and borrowers.

Finally, I use data on institutional ownership, the quality of external governance, board composition and macroeconomic conditions from Thomson Reuters Institutional (13f) Holdings, Institutional Shareholder Services (ISS), BoardEx, and Fred, respectively.

### 1.2.2 Firm and Loan Summary Statistics

Data from Compustat and DealScan show key features of breached firms: Large differences in total assets, profitability, capital and ownership structure, and stock volatility between firms.

Panel A of Table 1.2 displays firm characteristics for breached firms and the Compustat universe during the 2005-2015 period. Not surprisingly, target firms are bigger, more profitable, and make more substantial use of debt. This emphasizes the idea that larger firms may be more vulnerable to data breaches, while governance mechanisms (i.e. internal mechanisms such as oversight, auditing and employee's responsibilities) are more difficult to maintain, information technology investments or upgrades are more expensive and time consuming, and the value of stolen information is greater.

The average (median) breached firm has approximately \$18.8 billion (\$5.3 billion) of total assets, compared to just \$3.2 billion (\$382 million) for the Compustat universe. In addition, breached firms have higher profitability (before interests) and book leverage ratios than other firms. These figures hint at the mature nature of breached firms, which are often established, visible, and older than the average Compustat firm, which comprises of small, young and unprofitable entities. In fact, the average (median) age (in years) for target firms is 27.23 (22), compared to 19.58 (15) for non-breached firms, while book leverage ratio figures stand at 0.26 (0.19) versus 0.21 (0.15). Consistent with the established nature of target firms, stock volatility is significantly lower. $t$-tests for the equality in means reject the null hypothesis that breached and non-breached Compustat firms have similar average characteristics. Similarly, Wilcoxon Mann-Whitney tests reject the equality of medians.

Panel B of Table 1.2 presents mean, median, and $p$-values for the difference in means and medians of loan characteristics for target firms and for the DealScan universe. Table 1.2 shows that the mean (median) spread over LIBOR (in basis points) that target borrowers pay is 195.84 (162.5), compared to 222.27 (185) for the entire DealScan database. This difference is statistically significant at the 1% level. This clearly shows that breached firms receive more favorable loan terms from lenders than other companies, which reflects their better reputation, higher profitability, and likely stronger borrower-lender relationships.

Other descriptive statistics in Table 1.2 show that the mean (median) deal amount is \$1,325 million (\$750 million), versus only \$637 million for the DealScan universe (\$325 million). These figures are consistent with prior literature on bank loans.[4] On the other hand, both breached and non-breached borrowers rely on medium-term loans, with mean (median) for the two groups of 54.72

[4]See for example Campello and Gao (2017); Ivashina (2009); Nini et al. (2012)

(60), and 53.75 (60) months, respectively. Both differences in means and medians are not statistically different from zero.

The total number of covenants in a loan contract varies considerably between groups: Corroborating earlier discussions, loans to breached firms include fewer general covenants (i.e. have looser covenant intensity), than non-breached firms. The same conclusion applies to the presence of a financial covenant. Specifically, 58% of loans of breached firms include a financial covenant, while 62% for non-breached firms ($t$-test rejects equality in means at the 5% level). Similarly, loans of target firms are less likely to be secured by collateral (real estate, inventory, accounts receivable, etc.), compared to other firms, which signals risky borrowers who necessarily need to pledge more assets. I find no statistical differences between the two groups in terms of performance pricing provision.[5]

### 1.2.3 Research Design

To quantify the implication of a breach on a firm's external financing costs and debt contracting, I rely on a difference-in-difference specification on a matched sample (one to one match). I use propensity score matching on the natural logarithm of total assets, book leverage ratio, and stock volatility to identify similar non-breached firms.

I run the following model at the deal-level:

$$Y_{i,t} = \gamma_i + \lambda_t + \delta Breached_{i,t} + \Lambda X_{i,t} + \epsilon_{i,t} \tag{1.1}$$

where the dependent variable $Y_{it}$ represents the spread over the LIBOR, or a non-pricing feature of the loan. I run most of the specifications at the deal level, aggregating facilities that belong to the same package and keeping the largest (Campello and Gao (2017); Ivashina (2009)). Observations at the deal-level alleviate concerns that auto-correlation between facility of the same package reduces the standard errors. $\delta$ represents the difference in difference coefficient, and $X_{i,t}$ is a vector of controls. Loan controls include the natural logarithm of loan size (in $ million), the natural logarithm of loan maturity (in months), performance pricing indicator, financial covenant and secured dummy indicator. Firm controls include the natural logarithm of total assets, tangibility, profitability, cash

[5]Performance pricing consists of a grid displaying different pricing levels based on a predefined trigger such as a company's ratings, ratios, outstanding, etc.

flow volatility, stock price volatility, Altman z-score, book leverage ratio, Tobin's $q$, and firm's S&P rating. All specifications include loan type and deal purpose dummies as well. I cluster standard errors at the firm level to allow for within-firm correlation; however, significance of the results is robust to alternative specifications, such as independent-double clustering at the firm and year level. I define all the variables and computations in the Appendix.

Breached$_{i,t}$ takes a value of one for the three years following a data breach $(t+1, t+2, t+3)$ and the year of the breach $(t)$ whenever the breach event date is at least one month prior to the deal activation date, zero in the four years before. I include firm $(\gamma_i)$ fixed effects to control for time-invariant unobservable characteristics at the firm level, and year $(\lambda_t)$ fixed effects to control for time-varying market-wide shocks to firms. I also include loan type and deal purpose dummies to address concerns that loans for LBOs inherently differ from working capital loans. In robustness tests, I include industry-year fixed effects to account for time-varying industry-wide shocks. In fact, although data breaches occur in many industries, some are more susceptible than others to such attacks, and the changing nature of industries is likely to vary over time and to correlate with the probability of a data breach. This applies in particular to the retail and information technology sector, where more and more information is being stored in digital form today than a decade ago. This design allows me to rule out that industry-specific regulations or shocks drive the results.[6]

To confirm my results in a single difference framework, I follow Graham et al. (2008) and Chava et al. (2017). I run deal-level regressions for the sample of firms that experienced a date breach (treated). Specifically,

$$Y_{i,t} = \delta Post_{it} + FirmControls_{i,t} + LoanControls_{i,t} + MacroControls_t + \gamma_i + \epsilon_{i,t} \tag{1.3}$$

[6]I also look at the dynamic effects of data breaches through a difference-in-difference specification where I allow for leads and lags (pre- and post-treatment effects year by year) and:

$$Y_{i,t} = \gamma_i + \lambda_t + \sum_{\tau=-4}^{3} \delta_\tau Breached_{i,t+\tau} + \Lambda X_{i,t} + \epsilon_{i,t} \tag{1.2}$$

This model shows the average treatment effect the year of the breach, one, two, and three years after the event. Using a model with leads and lags, I can confidently include covariates that could be affected by the treatment (data breach) itself, without biasing the point estimates for the Breached$_{it}$ interaction, and alleviate concerns about violation of the parallel trend assumption.

where all dependent variables and controls are as in 1.1. Table 1.15 reports the results.

I also test for the cross-sectional treatment effects by type of data breached by running the following model:

$$Y_{i,t} = \gamma_i + \lambda_t + \delta \cdot \big(Breached_{i,t} \times Type\big) + \Lambda X_{i,t} + \epsilon_{i,t} \tag{1.4}$$

where breached type may refer to financial information loss, general privacy, employee related, or data breach through malicious outside entry. Finally, I test for mispricing and spillovers by running placebo tests in which I let either the *breached* firms, or the breach dates to be randomly assigned, and specifications based on *abnormal* loan spreads. Moreover, I run a loan-by-loan difference-in-difference where I found one control loan for each treated loan (see Section 1.6 for more information).

### 1.2.4 Identification

Any difference-in-difference empirical strategy relies on two main assumptions: (1) parallel (common) trend; and (2) *strict* exogeneity. Although direct tests of (1) and (2) are difficult (if not impossible) to carry out, I address the first by showing that breached and non-breached matched firms share statistically indistinguishable bank loan characteristics in the years preceding a data breach (see Table 1.4). Furthermore, Figure 1.5 reports the year-by-year difference-in-difference coefficients in a leads and lags controlled settings. Kolmogorov-Smirnov tests of equality of distribution also support the above conclusions (untabulated).

For the *strict* exogeneity condition not to hold, lenders (i.e. banks) must anticipate data breaches and change debt contracts in response (pricing and non-pricing contractual features alike), or firms must change their financial policies in anticipation of a breach. This is unlikely, unless disclosure of material events does not take place, and firms take remediation steps beforehand. Firm fixed effects control for any time-invariant unobservables correlated with the likelihood of a data breach, while year fixed effects control for common shocks across time (industry $\times$ year fixed effects control for a more subtle time-varying unobservable at the year-industry level).

### 1.2.5 Likelihood of Data Breaches

To understand which type of firm suffers from a data breach, I run probit regressions of the likelihood of experiencing a data breach each year on various firm characteristics and fixed effects:

$$\mathbb{P}(Breach_{it} = 1) = FirmControls_{it} + \lambda_t + \eta_j + \epsilon_{it} \tag{1.5}$$

The results echo those documented by previous research on cyberattacks (Akey et al. (2018); Kamiya et al. (2018)). Table 1.3 examines the determinants of data breaches.

The natural logarithm of total assets remains a good predictor of the probability of a data breach in all specifications. A one standard deviation increase in the natural logarithm of total assets is associated with a 68 basis points increase in the likelihood of a data breach. The magnitude is economically significant, given the unconditional probability of a first-time data breach of 0.3%. This is not surprising since larger and more established firms are more visible, thus constituting a valuable target. In addition, larger firms may be more susceptible to thefts of digital information, given the larger employee base and the geographic dispersion of offices and establishments across the country. Not surprisingly, less tangible firms (i.e. information technology), are more likely to suffer a data breach. A one standard deviation increase in asset tangibility is associated with a 17 basis points decrease in the probability of a data breach. In contrast to common belief, breached firms are more profitable, reinforcing the view that targeted firms are mature but otherwise healthy and profitable firms. With a marginal coefficient of 0.014, the effect remains strong. Moreover, stock volatility, measured by the standard deviation of daily stock returns over the previous fiscal year, correlates negatively with the likelihood of a data breach. Breached firms also have a higher Tobin's $q$.

Columns 4 to 6 add additional controls, such as the natural logarithm of firm's age (in years), cash flow volatility, and the percentage of share held by institutional investors. Notably, a higher institutional ownership predicts a larger probability of a data breach.

## 1.3 Main Results

### 1.3.1 Loan Spread Regressions

This section describes the baseline results of Equation 1.1. I find one match (control) for each breached firm (treated) using propensity score matching on the natural logarithm of total assets, book leverage ratio, and stock market volatility.[7] I also require treated and control firms to be in the same fiscal year and industry (in the same 2-digit Standard Industry Classification). I assign a

[7]Results are qualitatively and quantitatively similar when I match on different covariates, such as profitability, Tobin's $q$, etc.

synthetic event date to each control firm on the basis of the actual breach date of its paired treated firm. I use four years of data before and four years after an event date, although the sample can be restricted to a narrower window without substantially changing the results.

Panel A of Table 1.4 shows firm characteristics the year prior to a data breach. The matching algorithm properly identifies similar firms along several dimensions. Average (median) total assets stand at $15.5 billion ($4.1 billion), and $11.1 billion ($3.3 billion) for treated and control firms, respectively. Tests for both the means and the medians suggest that these are not statistically different from each other. The two groups also behave similarly along the capital structure, profitability, and governance dimensions. Figure 1.4 in the Appendix shows that the kernel density of treated firms is similar to those of control firms and Kolmogorov-Smirnov tests of equality of distribution support this conclusion.

Looking at Panel B of Table 1.4, the average (median) breached firm pays 179.41 (172.50) basis points on its bank loans, versus 187.01 (150) for control firms in the four years preceding a data breach. These differences are statistically indistinguishable from zero. Moreover, if I extend the sample to any year prior to a data breach, spreads paid by the two groups are still statistically indistinguishable from each other, indicating that the effect (if any) of higher spreads is specific to the post-breach period. Similarly, loan maturity averages 4 to 5 years (medium-term notes) for both treated and control firms, while the average loan size (in millions of dollars) is $1,005 and $904, respectively. Debt contract features and covenants do not differ as well.

Table 1.5 reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of a data breach on loan spreads. The estimation follows Equation 1.1. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR (or equivalent) in basis points. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event, including the year of the breach.

Column (1) of Table 1.5 shows that the difference-in-difference coefficient is 0.18, while only controlling for loan maturity and loan size. The effect is significant at the 1% statistical level. Economically, the magnitude is notable and translates to a 20% increase in the interest rate paid by borrowers over LIBOR. Consistent with prior studies on bank loans, the natural logarithm of the loan amount is negatively associated with loan spreads, while loan maturity is not significantly correlated at conventional levels.

Columns (2) to (5) repeat the estimation and add more loan and firm characteristics to the baseline specification to control for other important determinants of loan spreads. Column (2) adds debt contractual features such as collateral use, financial covenant and performance pricing provision. As expected, firms pledging a collateral are associated with higher spreads, while neither financial covenant nor performance pricing provisions are statistically different from zero. Column (3) includes firm's rating, as provided by Standard and Poor's. I also include a dummy that takes a value of one if the firm is not rated, zero otherwise. The rating dummy can then take an arbitrary number for the missing category. As expected, the higher the rating, the lower the spread a borrower pays. This is as expected, as bankers frequently use credit ratings as the basis for adding a risk-based spread.[8]

Columns (4) and (5) add macroeconomic variables (credit spread and term spreads) measured the month preceding the deal activation date, and firm specific controls. Supported by the discussion about the matching procedure above, most coefficients are not statistically significant. Note that leverage would predict a positive coefficient. However, I find a negative coefficient, which is not surprising given that breached firms are slightly more levered than control firms (although not significantly so). However, across all specifications, the difference-in-difference coefficients remain large and statistically significant. Furthermore, restricting the sample to revolvers and lines of credit with more than 365 days of maturity leaves all results unchanged (not tabulated). Similarly, focusing on loan for working capital and corporate purposes also leads to identical conclusions (not tabulated). This suggests that low-quality loans or loans for aggressive business purposes (takeovers, LBOs) do not drive the results.

I also replicate some of the earlier studies in the literature (e.g. Graham et al. (2008)) and document an economically significant effect of data breaches on loan spreads for treated firms. On the other hand, control firms do not experience a significant change in the spread. Table 1.15 in the Appendix summarizes these results.

To summarize, compared to private loans initiated by similar firms before and after a data breach, *breached* firms pay larger spreads. The increase ranges from 15 to 20 percent. Given an unconditional

[8]Results are qualitatively and quantitatively similar if I include separate dummies for each ratings, or split the sample based on investment grade ratings vs high yield status

average of 179 basis points in the four years prior to a data breach, this translates into a 30 to 40 basis points increase in the interest rate charged by lenders. In dollar terms, this averages to 3 to 5 million dollars in additional interests paid by borrowers each year *per* loan, which translates to $17 million over the average life of the loan. Given an average 12 month LIBOR of 2.15 percent over the sample period, the documented increase in spread after a data breach translates into a 8 to 10 percent increase in the total loan interest rate. I extrapolate these results to the overall balance sheet of *breached* firms, and come up with an estimate of the overall loss in value for debt holders. If all debt was traded, and the change in interest rate was constant across debt types and maturities, the overall percentage change in debt value would range from 1.25 to 1.75 percent, which is economically as large as the stock market drops documented in the literature.[9] Overall, these effects are comparable in magnitude to those documented by Deng et al. (2014) in the context of shareholders' litigation, Graham et al. (2008) and Chava et al. (2017) for corporate misreporting, and Yuan and Zhang (2015) in documenting the effect of class action lawsuits on loan prices. However, only a handful of companies disclose data breaches, since firms realize that cyberrisk threatens customers, shareholders, and creditors. Moreover, some firms might find it prohibitively expensive to refinance their loans or tap the debt capital market, hence avoiding it altogether. Therefore, my findings likely represent a lower bound for the true cost of data breaches on debt financing.

### 1.3.2 Heterogeneous Treatment by Breached Information

According to a security report by Verizon (2018), 76% of breaches were financially motivated (https://enterprise.verizon.com/resources/reports/dbir/). The potential tangible losses stemming from stolen financial information such as credit and debt card, bank account number and password related credentials, range from spending toward credit monitoring and protection for customer, legal costs associated with lawsuits, as well as other form of disbursement. I therefore hypothesize that lenders price the effect of financial information loss more than other type of losses (customer or

---

[9] I compute approximate values of bond durations for each firm using a weighted average of future debt maturities in the year preceding a data breach. I found average durations ranging from 3.15 to 4.35 years. I apply the standard duration formula:

$$\frac{\Delta P}{P} = -D_{mod} \cdot \Delta r$$

where $D_{mod}$ is the modified duration, and $\Delta r$ is the change in spread in basis points after a data breach.

employee related, general privacy). Table 1.6 studies the heterogeneous average treatment effect by type of breached information. I focus on the loss of financial data (credit and debit card, bank account number, and credentials to access a financial platforms or account) versus loss of general data (privacy data, such as name, address e-mail, etc.) or employee information. I also estimate the differential effect of malware third-party entry.

Column (1) shows that compared to non-financially motivated breaches, financially motivated breaches increase the spread by 20 to 30% more compared to the baseline difference-in-difference coefficient. This indicates that the threat of repercussion on firms health and future expected cash flows is larger when financial (and hence customer) related information is involved.

When I consider other types of stolen information, I do not find significant differences. Columns (3) and (4) replicate the previous two columns for breaches of employee data (e-mails, names, addresses, etc.), while column (4) restricts the sample to general privacy data loss (non-financial). The effect is still positive and of economically significant magnitude. Interestingly, column (2) shows that first-time malware third-party entries lead to a three times larger impact on bank loans, while the baseline interaction is still economically and statistically significant. Column (5) shows that the effect is more pronounced for larger data breaches, i.e. where more records are stolen.

Overall, these results indicate that creditors respond to financially motivated breaches by charging higher spreads. The effect is economically large, and likely a function of future expected cash flows and their risk. Banks understand that litigation risk and remediation costs are likely larger for the subsets of borrowers that experience these types of breaches.

## 1.4 Debt Contractual Features, Covenants and Other Terms

Bank loans contain several features, specified by the contract between the borrowing firm and the lending institution. These range from the maturity and amount outstanding, to the use of general and/or financial covenants or other provisions. I study the effect of data breaches on other pricing and non-pricing features of the loan agreements. If lenders perceive cyberrisk as material and conveying hard-to-process information about future cash flows, litigation risk, increased management inattention going forward, and increased government scrutiny, they may respond by changing debt contractual features *and* interest rates.

Panel A of Table 1.7 analyzes the effect of data breaches on other pricing and non-pricing terms of the loan contract, such as the total loan amount, maturity, the number of lenders in the syndicate, and the fees charged by the lead arranger.

Columns (1) and (2) show that borrowers do not alter the dollar amount they borrow, nor they change the maturity structure of their debt. Although I cannot conclude whether demand or supply (or both) shifts, there seems to be little to no effect on the maturity profile of bank loan agreements. On the other hand, the borrowed amount per loan likely increases. This is particularly true in the single-difference setting of Table 1.15.

The size of the syndicate plays an important monitoring role. For example, a more concentrated syndicate with fewer lenders find it easier and less costly to control borrowing firms and share information among themselves. The type of risk posed by data breaches may cause changes in the syndicate structure. However, I do not find evidence of this.

On the pricing side, lenders usually charge fees in the form of *annual* and *upfront* fees. While annual fees are charged annually against the entire commitment value, whether used of unused, upfront fees are paid by the borrower to the lead arranger at the closing of a loan. The lead arranger decides to share the fee with other participants in the syndicate, if deemed necessary. Columns (4) and (5) show that part of the increase in loan spread comes from an increase in the annual fee charged, as opposed to an increase in the upfront fee.

Panel B of Table 1.7 studies how lenders react to data breaches by modifying the contractual features agreed upon signing a loan. Column (1) shows that covenant intensity, as measured by the natural logarithm of one plus the total number of general covenants, increases by approximately 30%. Therefore, banks respond to a negative shock in their lending portfolio by including additional covenants that restrict the ability of a firm to make poor investment decisions which decrease the likelihood of debt repayment going forward. This shifts part of the control power to creditors, mitigating misbehavior by managers. These provisions include cash and debt sweeps, equity sweeps, and dividend restrictions. For example, lenders require firms to repay part of their bank loans with a portion of excess cash flows, excess net long-term debt, or equity. In other words, data breaches alter the expected riskiness of future cash flows, for example through an increase in volatility. Cash sweep provisions create an additional buffer in the event of future adverse shocks to a firm's stream of cash flows. In the context of project finance deals, Corielli et al. (2010) find that lenders rely on the network of non-financial contracts as a mechanism to control agency costs and project risks.

Column (2) replaces the total number of covenants with a dummy variable that takes a value of one if the terms of the contract include at least one general covenant, zero otherwise. The magnitude and significance of the coefficient confirms the results of column (1), i.e. borrowing firms are 14% more likely to have a general covenant included in their contracts.

Column (3) studies the use of performance pricing provisions and finds no significant relationship, while column (4) shows that there is a statistical significant change in the use of financial covenants at the 10 percent level (accounting and performance ratios). In addition, lenders do not require borrowing firms to pledge additional collateral in order to secure the loan.

To understand where the change in debt contracting comes from, I retrieve information about the specific type of covenants included, such as asset, debt, and equity sweep, dividend restrictions, and use of excess cash flows. Column (1) through (5) of Table 1.8 report the results. Overall, lenders are more likely to include debt sweep and dividend restriction provisions. This goes hand in hand with the high-dividend paying nature of breached firms, as well as their larger leverage base vis-à-vis smaller and younger firms.

Overall, results of Panel A indicate that lenders respond to data breaches to their portfolio companies by tightening (general) covenant intensity. This is consistent with the view that general covenants are easier to implement and facilitate the monitoring role by banks (Graham et al. (2008)).

## 1.5 Cross-Sectional Analysis

I examine whether there are cross-sectional differences in response to data breaches based on *ex*-ante firm characteristics, syndication, and contract features. I divide companies into terciles based on observable characteristics at time $t-1$ and estimate 1.1 on the samples of firms in the upper and lower terciles.

Panel A of Table 1.9 divides firms based on measures of visibility (as proxied by the Fortune 500 indicator), profitability, and capital investments. I find that lenders significantly increase the interest rate when the borrowing firm is a Fortune 500. On the other hand, I find no effect for firms outside the ranking. Lower profitability firms tend to be charged higher spreads than their higher profitability counterparts; however, the two coefficients are not statistically different from each other at conventional levels. Interestingly, lenders charge capital intensive firm higher spreads. Overall, these results suggest that lenders view data breaches as negative shocks to large and visible firms, with lower profitability, and requiring large capital expenditures. It may be the case that lenders realize that IT updates and cyber-security investments are costly, and may drive performance and profitability lower.

Panel B of Table 1.9 considers firm's characteristics such as the presence of a cyberrisk officer, a measure of bankruptcy risk, and whether a firm pays high or low dividends. I look for the presence of a cyberrisk officer in the year(s) preceding a data breach for each treated and control firm, using data from BoardEx. I find that firms that do not have a cyber risk officer before the event experience a greater increase in spreads. The coefficient for the interaction term is statistically significant at the 5% level, and more than double in magnitude. I also find that the effects of larger spreads come from low bankruptcy risk firm (high Altman Z-score), which is inconsistent with the view that *ex*-ante highly distressed firms suffer more from data breaches. The last two columns show that the effect on high dividend firms is more pronounced. This is consistent with the findings that lenders adjust the debt contracts by including provisions that restrict dividend payments and with dividends being a indirect measure of financial constraint.

Another important dimensions to study is the relationship between the interest rate charged, and the composition of the syndicate during the lending activity. The syndication process depends on complex interactions between a lead arranger (or multiple lead arrangers), and the syndicate participants. During the syndication, the lead arrangers decide what percentage of the loan to retain for themselves and how much to allocate to other participants. The lead arrangers will hold a portion $\alpha$ of the loan and will set a required spread over the base rate. At the same time, lead participants demand a spread based on the signal ($\alpha$) they receive from the lead arrangers. As documented by Ivashina (2009), the effect of a larger share $\alpha$ retained by the lead arranger implies both a *decrease* in asymmetric information, *and* an *increase* in the portfolio (idiosyncratic) credit risk of the lead arranger. On the other hand, a decrease in $\alpha$ leads to lower diversification risk but larger moral hazard and adverse selection.

Table 1.10 displays results based on the characteristics of the syndicate. I construct the average share $\alpha$ retained by the lead arranger for the deals in the four years preceding a data breach, as well a Herfindahl-Hirschman Index of ownership concentration among participants as commonly done in the literature (see for example Sufi (2007)).

Column (1) through (3) all point at the same conclusion: The increase in spread comes from firms whose *ex*-ante loan share retained by the lead arranger is high, when there is less concentration in lenders, and fewer of them. This is more consistent with a shift in the idiosyncratic credit risk of the lead arranger's portfolio. These firms are likely to be highly monitored, where lenders decreased the asymmetric information by retaining a larger share of the loan. After a breach, a likely increase in the portfolio credit risk, coupled with a decrease in the share retained (unreported results show an economically significant decrease in the average share retained by the lead after a data breach, although not statistically significant at conventional levels) might have lead to higher spreads charged by participating banks.

### 1.6 Channels and Alternative Explanations

Why do banks alter their lending policy after a data breach, by charging higher spreads and intensifying their monitoring through tighter debt covenants? I explore channels that may corroborate the empirical findings or reject alternative explanations. Two non-mutually exclusive explanations exist. First, in the aftermath of a data breach, investors and lenders realize that breached firms were *"lemon"*, i.e. data breaches are more likely to happen at weaker firms with deteriorating fundamentals in place, revealing their type as *"bad"*. Second, cyberrisk poses a true risk to a firm's future operations through changes in performance and capital metrics which banks strictly monitor.

In light of these alternative explanations, I first study the effect of data breaches on firms' outcomes. In particular, return on assets decreases (net income to total assets), and financial and operating leverage increases. Reasons may include temporary loss of customers, increase litigation risk, lump investments in IT infrastructure and risk management practices (Kamiya et al. (2018)), actions to rebuild reputation (increase spending in extraordinary charges and CSR as in Akey et al. (2018)), and high level of management distraction going forward. Additionally, investors may reassess the probability of future negative events, which makes rare risky episodes more likely and cyber crisis more dangerous. In fact, I find that distress probability increases, as measured by the Merton's distance to default. Moreover, for the subset of loans that contain financial covenants, I link changes in profitability and capital structure to borrowers' covenant violations and banks' behavior. I compute measures of probability of covenant violation at inception (see Demerjian and Owens (2016) for details on how to compute various measures of covenant violation), and show that the effect is partly driven by borrowers more likely to breach a capital rather than a performance covenant.

On the other hand, lenders may misprice loans in the first place, by charging lower spreads and contracting on looser covenants. I alleviate concerns about mispricing in three ways: First, I run *abnormal* spread (and covenant) regressions and show that differences persist after a data breach when correctly pricing loans based on observable characteristics; second, I repeat the main specification of Equation 1.1 (and the abnormal spread regression discussed above) on a loan-by-loan difference-in-difference basis; and lastly, I perform placebo tests with respect to both the timing of a breach and the treatment firms (or loans).

### 1.6.1 Changes in Fundamentals

**Firm Outcomes** Table 1.11 studies the effects of data breaches on firm outcomes. This table finds similar economic magnitudes to those documented by Kamiya et al. (2018), the *first* to shed light on the effect of cyber-attacks on target firms. The findings suggest that firms change their capital structure (financial and operating), and profitability (as measured by ROA) decreases. I do not find significant changes in the ratio of EBIT and EBITDA to total assets. This suggests that higher interests and/or extraordinary items drive the decline in profitability. Moreover, distress probability increases. I also find that the ratio of operating leases to total assets increases, consistent with investments in software and gears via leasing agreements and debt issues.[10]

**Likelihood of Additional Data Breaches** Table 1.12 studies the likelihood of a future data breach, conditional on a breach occurring in year $t$ or $t-1$. Therefore, I restrict the sample to include all firms breached at least once over the sample period. $\text{Breach}_t$ takes a value of one if a breach happens in year $t$, zero otherwise. Similarly, $\text{breach}_{t-1}$ equals one if a breach happens in year $t-1$, zero otherwise. I find a positive and significant relationships across most specifications, which suggests that, conditional on a data breach event, the likelihood of a second breach increases.

Columns (1) through (3) constraint the sample to firms that experienced at least one data breach over the sample period. Columns (2) and (3) use the contemporaneous and first lag of the data breach indicator as an independent variable. The marginal coefficients correspond to a 4 to 6 percent increase in the likelihood of a second data breach. Moreover, the probability of a second data breach increases in the first subsequent year but drops after 2 or more years.

Columns (4) to (6) restrict the sample to firms with at least two data breach events over the

[10] I compute the present value of minimum rental commitments as in Rauh and Sufi (2011) using a 10 percent discount rate

sample period. The marginal coefficients correspond to a 10 to 12 percent increase in the likelihood of a second data breach in the following year.

Overall, results of Table 1.12 suggest that investors reassess upward the likelihood of negative shocks to reputation. The effect is particularly strong in the first year following a data breach, hinting at the wave-like nature of breaches. Although the unconditional probability of a data breach is small, the increased likelihood of an additional data breach suggests a change in the higher moments of the distribution of events, such as a thickening of the tails.

**Probability of Covenant Violations After a Data Breach** Relationship lending relies on continuous interactions between lenders and borrowers, and frequent monitoring of good financial and operating standing. This is especially true for loans with covenants, the breach of which would lead to technical default and increased scrutiny. Therefore, any causal changes in performance or capital structure may lead to increased probability of covenant violation – ultimately translating into higher interest rates. To test whether lenders charge borrowers higher interest rates as a function of their probability of covenant violation at inception after a data breach, I interact $breached_{it}$ with a measure developed by Demerjian and Owens (2016).

Table 1.13 shows that the increased interest rate comes from borrowers that are more likely to violate capital covenants rather than performance-based ones. Columns (5) and (6) show that, even after controlling for the baseline probability of covenant violation at contract inception, breached firms with higher probability of capital covenant violation pay larger spreads than breached firms with lower probability of capital covenant violation.

### 1.6.2 Mispricing by Lenders

**Abnormal Spread Regressions** I have documented so far that lenders modify loan pricing and non-pricing features following a data breach. However, it may be the case that breached firms are *"mispriced"* by banks. For instance, firms that have never experienced a data breach and have enjoyed superior growth over the years may receive favorable terms from lenders. Consequently, the effect of a data breach simply translates to an upward price adjustment to levels consistent with each firm's risk-profile and loan characteristics. To address this hypothesis, I conduct a two-stage analysis. In the first pass, I estimate a loan spread regression on the entire sample of firms, excluding entities that have been breached. If banks consistently misprice borrowers, loan and firm characteristics should explain the price paid over LIBOR, and I should see no difference in spreads before and after a data breach. Therefore, I estimate the price that breached firms would pay as the estimated fitted value. The residuals $\hat{\epsilon}_{i,t}$ (actual minus fitted spread charged) represent the abnormal spreads. If banks consistently misprice, I should not observe any significant effect after the adjustment. Column (2) of Table 1.14 shows that the effect of a data breach is still sizable and of similar magnitude (similarly for covenant intensity, i.e. banks did not consistently arrange loose debt contracts, on average). Overall, mispricing of risk by lenders does not seem to drive my results.

**Loan-by-Loan Analysis** I perform a loan-by-loan difference-in-difference test which allows me to rule out the idea that consistent unobserved differences between firms (breached and controls) may account for the effects I document. Specifically, for each loan by a breached firm, I find a similar loan, based on firm and loan characteristics. Therefore, a treated firm's loan can be matched to loans belonging to different firms. As before, I require firms to be in the same 2-digits SIC, same year, and same loan type and deal purpose. In addition to the matching covariates used in the baseline specification (size, leverage, stock volatility), I require loans to be of similar size and maturity. Columns (4) and (5) of Table 1.13 present baseline and abnormal spread results. Again, magnitudes and statistical significance are similar to before, indicating that consistent mispricing does not fully explain my results.

**Placebo Tests** To alleviate concerns that an omitted variable drives the increase in loan spread for breached firms, I perform two sets of placebo tests in which I either use a different treatment group, or a different event date for the *actual* original group. In other words, the first test assumes that non-breached (now treated) firms should behave in the same way as non-breached control firms. The second test assumes that the treatment effect should be specific to the actual treatment period (see Almeida et al. (2012)). In the first test, I create 1,000 random samples of "pseudo-breached" firms (or loans) with random assignment of *actual* event dates. I then repeat the analysis of Section 1.3 and compute difference-in-difference coefficients and $t$-statistics. Figure 1.2 plots the result. Both the average coefficient and $t$-statistics are centered around zero, with a median $\beta$ of 0.008 and $t$-statistic of 0.05. Less than 1 percent of placebo coefficients are statistically significant at conventional levels.

The second test uses the *original* sample of treated firms (loans) but assigns a *synthetic* random event date prior to the actual date. The right panel of Figure 1.2 shows the approximate normality of the distribution of both the difference-in-difference coefficients and $t$-statistics. As before, median values are 0.005 and 0.05, respectively. Overall, breach events appear to be real occurrences, both in terms of their timing *and* the target firms. Columns (7) and (8) repeat the two placebo tests by drawing 1,000 random loans (not firms), and dates. Again, results are consistent with the analysis done at the firm level.

## 1.7 Robustness

Although the analysis in Section 1.3 suggests that *breached* firms pay larger spreads on bank loans than similar *non-breached* firms after a data breach, I carry out a battery of robustness tests to strengthen the validity of my results. First, I add more controls and fixed effects to the empirical model of Regression 1.3. Second, results hold true when: (1) I match on different observable characteristics; (2) I use more than one control firm; (3) I consider repeated breaches to the same firm.

**Additional Controls.** Table 1.19 runs a battery of robustness tests by including additional control variables that affect both the composition of the sample and the coefficient of interest (when I include ownership and governance variables I lose about one-fifth of the observations). However, the difference-in-difference estimates remain economically and statistically significant across all specifications, and of similar magnitude to the baseline results of Section 1.3. Column (1) controls for the percentage of shares held by institutional investors and finds similar coefficients to the one reported in Section 1.3. Columns (2) to (4) control for widely used measures of financial constraint. In particular, debt constraints usually reflect existing leverage, or covenant violations rather than information frictions. Column (2) uses the Kaplan-Zingales Index, column (3) the Whited-Wu Index and column (4) proxies constraints with the natural logarithm of age (in years). Older firms may be more financially constrained, highly levered, and thus age may explain the cross-sectional variation in spreads. On the other side, younger firms may have more pronounced information frictions and syndicate members require premia and/or tighter covenants. Although these measures are not widely included in the empirical literature on bank loans, I follow Chava et al. (2017) and show the robustness of my results.

Column (5) independently double-clusters standard errors at the firm and year level. Column (6) includes a treatment-specific linear time trend. The coefficient remains of similar magnitude, which indicates that both treatment and control groups were not on a differential path prior to a breach. Column (8) adds industry $\times$ year fixed effects to control for unobservable time-varying industry shocks. Finally, columns (9) and (10) remove loan type and purpose fixed effects, and year fixed effects, respectively.

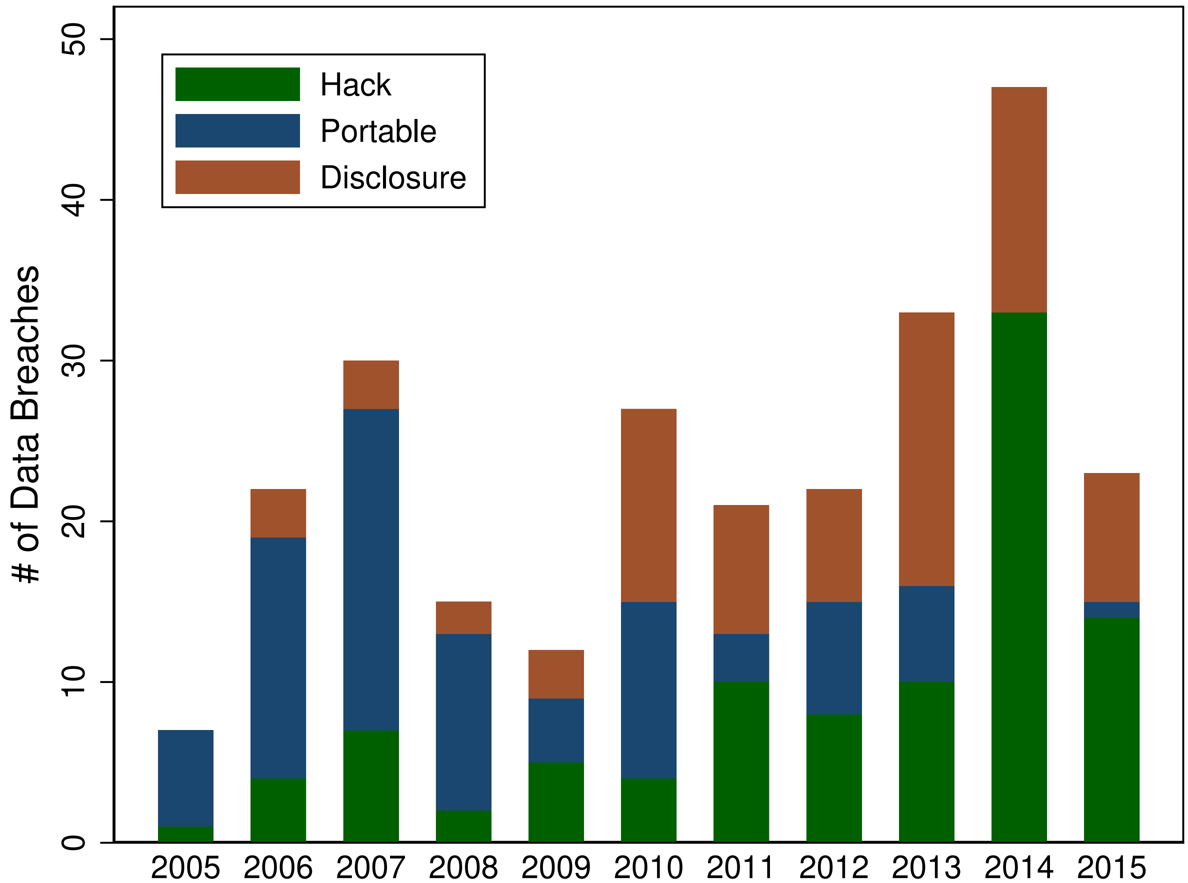
**Matching.** I repeat the analysis of Section 1.3 by performing a one-to-two propensity score matching. Results are qualitatively and quantitatively similar. Matching on different covariates such as profitability, Tobin's $q$ or institutional ownership leads to comparable results. Finally, using repeated data breaches to the same firm leads to similar results.

## 1.8 Conclusions

What are the *real* and *financial* costs of data breaches? I document an economically important effect of data breaches on firms' financing costs and debt contracting. I find that breached firms pay 20% larger spread than similar non-breached firms after an event. Lenders tighten covenant intensity, consistent with a shifts in control rights over cash flows. Banks respond more aggressively to breaches of financial information, rather than general privacy data. Furthermore, the effect is stronger for capital intensive, visible firms, but with a lack of cyber security measures in place. Consistent with previous literature, profitability declines and financial leverage increases. Firms also make more extensive use of operating leases, consistent with investments in software and IT equipment. Data breaches convey information about risk management vulnerabilities and leads to increased default risk, to which banks respond. My findings are economically significant, and likely represent a lower bound for the effect of data breaches on firm's financing costs. In dollar terms, borrowers pay 3 to 5 million dollars in additional interests each year per loan, which translate to $17 million over the average life of the loan. These findings are of similar magnitude to those documented in the context of shareholders' litigation and corporate financial misreporting.

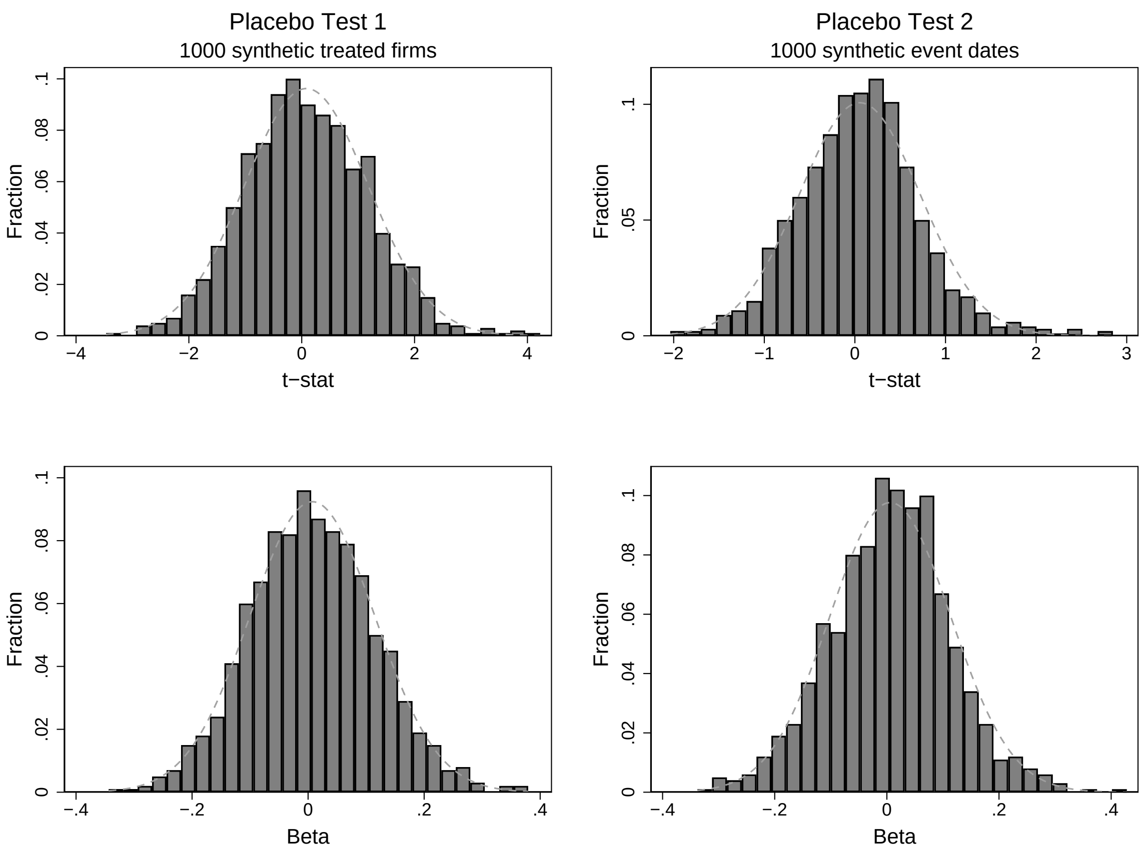
While the level of attention and scrutiny paid by firms and the public to data breaches have increased dramatically over the last decade, security infrastructures cost companies millions of dollars each year, but young firms often prioritize other aspects of their businesses, such as increasing production, research and development, and recruiting. In many cases, companies are unprepared and vulnerable, making the potential marginal investment in information technology and cybersecurity very large. Similarly, mature companies with a large employee base may find it difficult to successfully train employees on data-privacy issues, making the entire system fragile. The investment and labor market implications of cyberrisk represent interesting avenues for future research.

**Figure 1.1: Number of Data Breaches by Type**




This figure shows the total number of data breaches by type for US publicly listed firms between 2005 and 2015. The figure depicts the three most common types of data breach, hacking, portable device, and disclosure. Hacking refers to attacks by an outside party (i.e. cyber-attacks), portable device refers to lost or stolen physical drive containing digital information, and disclosure refers to unintended disclosure of sensitive digital information.

**Figure 1.2: Robustness Test: Placebo Regressions**




These figures report average OLS regression coefficients and $t$-statistics from 1,000 placebo deal-level difference-in-difference regressions for the effect of data breaches on the cost of bank loans. The first test consists of 122 *pseudo-breached* firms with synthetic event dates, randomly chosen from the Compustat universe. The second test consists of the actual 122 *breached* firms, with synthetic event dates before the actual data breach. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points.

**Table 1.1: Summary Statistics for Data Breaches**

| | N | Mean | SD | Median |
|---|---|---|---|---|
| | | Panel A: Type of Breach | | |
| CARD | 4,880 | 0.01 | 0.11 | 0.00 |
| DISC | 4,880 | 0.16 | 0.37 | 0.00 |
| HACK | 4,880 | 0.23 | 0.42 | 0.00 |
| INSD | 4,880 | 0.10 | 0.30 | 0.00 |
| PHYS | 4,880 | 0.26 | 0.44 | 0.00 |
| PORT | 4,880 | 0.16 | 0.37 | 0.00 |
| STAT | 4,880 | 0.03 | 0.18 | 0.00 |
| UNKN | 4,880 | 0.04 | 0.20 | 0.00 |
| | | Panel B: Type of Breached Information | | |
| Financial | 4,880 | 0.21 | 0.41 | 0.00 |
| Medical | 4,880 | 0.38 | 0.49 | 0.00 |
| Privacy | 4,880 | 0.62 | 0.49 | 1.00 |
| Employee | 4,880 | 0.26 | 0.44 | 0.00 |
| Total Records (1,000s) | 4,880 | 640.94 | 15,700.51 | 0.80 |

This table reports summary statistics for data breaches over the period 2005-2015 from the Privacy Rights Clearinghouse's Chronology of Data Breaches (PRC). The table shows the number of data breaches (N), mean, standard deviation, and median values. PRC provides the following definitions of data breaches: CARD includes fraud (credit and debit card), not via hacking; DISC refers to unintended disclosure (not involving hacking, intentional breach or physical loss); HACK refers to hacking by a third party; INSD is someone with legitimate access intentionally breaches information - such as an employee, contractor or customer; PHYS includes paper documents that are lost, discarded or stolen (non electronic); PORT refers to lost, discarded or stolen laptop, PDA, smartphone, memory stick, CDs, hard drive, data tape, etc.; STAT refers to stationary computer loss (lost, inappropriately accessed, discarded or stolen computer or server not designed for mobility). *Financial* refers to breach of credit and debit card information, and bank account numbers or credentials to access online banking platforms; *medical* is the loss of medical and patients related data; *privacy* refers to general privacy data such as SSN, e-mails, names, addresses, etc.; *employee* refers to employee's data; *total records* is the total number of records breached.

Table 1.2: **Summary Statistics for Breached and Non-Breached Firms**

| | Panel A: Firm Characteristics | | | | | |
|---|---|---|---|---|---|---|
| | Breached (A) | | COMPUSTAT (B) | | Difference (A-B) | |
| | (N = 1,176) | | (N = 34,091) | | (Mean) | (Median) |
| | Mean | Median | Mean | Median | $p$-value | $p$-value |
| Total Assets ($ million) | 18,845.83 | 5,335.44 | 3,251.72 | 381.96 | 0.00 | 0.00 |
| Tangibility | 0.22 | 0.13 | 0.24 | 0.15 | 0.00 | 0.81 |
| Profitability | 0.10 | 0.09 | -0.03 | 0.06 | 0.00 | 0.00 |
| Sales Growth | 0.10 | 0.07 | 0.88 | 0.06 | 0.71 | 0.00 |
| Leverage | 0.26 | 0.19 | 0.21 | 0.15 | 0.00 | 0.00 |
| Tobin's $q$ | 2.13 | 1.67 | 2.38 | 1.57 | 0.75 | 0.00 |
| CAPX/Assets | 0.04 | 0.03 | 0.05 | 0.03 | 0.00 | 0.02 |
| Cash Flow Volatility | 1.91 | 1.30 | 0.93 | 0.51 | 0.00 | 0.00 |
| Stock Volatility | 0.37 | 0.32 | 0.55 | 0.46 | 0.00 | 0.00 |
| Inst. Ownership | 0.76 | 0.79 | 0.58 | 0.64 | 0.00 | 0.00 |
| | Panel B: Deal (Loan) Characteristics | | | | | |
| | Breached (A) | | DealScan (B) | | Difference (A-B) | |
| | (N = 703) | | (N = 11,098) | | (Mean) | (Median) |
| | Mean | Median | Mean | Median | $p$-value | $p$-value |
| Spread (bps) | 195.84 | 162.50 | 222.27 | 185.00 | 0.00 | 0.00 |
| Maturity (months) | 54.72 | 60.00 | 53.65 | 60.00 | 0.14 | 0.36 |
| Amount ($ million) | 1,324.96 | 750.00 | 637.23 | 325.00 | 0.00 | 0.00 |
| Covenants | 1.20 | 1.00 | 1.65 | 1.00 | 0.00 | 0.00 |
| Financial Covenant | 0.58 | 1.00 | 0.62 | 1.00 | 0.03 | 0.00 |
| Performance Pricing | 0.47 | 0.00 | 0.45 | 0.00 | 0.22 | 0.63 |
| Secured | 0.39 | 0.00 | 0.55 | 1.00 | 0.00 | 0.00 |

This table reports summary statistics for firms that experienced a data breach (*breached*) and the Compustat universe over the period 2005-2015. Panel A reports mean and median values for *breached* and Compustat firm characteristics, as well as $p$-values for the test of difference in means and medians. Firm's characteristics are measured over the entire sample period. Panel B reports mean and median values for loans of *breached* and DealScan firms, and $p$-values for the test of difference in means and medians. Mean differences are measured using the $t$-test; median differences are tested using the Wilcoxon Mann-Whitney test.

**Table 1.3: Determinants of Data Breaches**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Total Assets | 0.22*** | 0.22*** | 0.26*** | 0.22*** | 0.22*** | 0.24*** |
| | (0.02) | (0.02) | (0.02) | (0.03) | (0.03) | (0.03) |
| Tangibility | -0.40** | -0.40** | -0.34 | -0.47* | -0.47* | -0.43 |
| | (0.20) | (0.20) | (0.24) | (0.24) | (0.25) | (0.28) |
| Profitability | 1.05*** | 1.07*** | 0.98*** | 0.99*** | 0.98*** | 1.02*** |
| | (0.25) | (0.25) | (0.33) | (0.28) | (0.28) | (0.37) |
| Sales Growth | -0.01 | -0.01 | -0.00 | -0.00 | -0.00 | -0.00 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | (0.00) |
| Leverage | -0.00 | -0.00 | -0.06 | -0.03 | -0.03 | -0.06 |
| | (0.18) | (0.19) | (0.20) | (0.22) | (0.22) | (0.22) |
| Tobin's $q$ | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| CAPX/Assets | 0.64 | 0.58 | 0.60 | 0.45 | 0.43 | 0.69 |
| | (0.66) | (0.67) | (0.79) | (0.93) | (0.95) | (1.04) |
| Stock Volatility | -0.27* | -0.19 | -0.13 | -0.45*** | -0.50** | -0.44* |
| | (0.14) | (0.19) | (0.20) | (0.16) | (0.23) | (0.23) |
| Age | | | | -0.06 | -0.07 | -0.02 |
| | | | | (0.04) | (0.05) | (0.05) |
| Cash Flow Volatility | | | | 0.04 | 0.04 | 0.04 |
| | | | | (0.04) | (0.04) | (0.04) |
| Inst. Ownership | | | | 0.43*** | 0.44*** | 0.29 |
| | | | | (0.16) | (0.17) | (0.21) |
| Year Fixed Effects | No | Yes | Yes | No | Yes | Yes |
| Industry Fixed Effects | No | No | Yes | No | No | Yes |
| Pseudo-$R^2$ | 0.15 | 0.16 | 0.20 | 0.15 | 0.17 | 0.21 |
| Observations | 40159 | 40159 | 40159 | 33304 | 33304 | 33304 |

This table reports coefficients and standard errors from probit regressions for the likelihood of a data breach on firm's characteristics. The sample consists of firm-year observations from Compustat over the period 2005-2015. The dependent variable is a dummy variable that takes a value of one if a firm experiences a data breach in a given year, zero otherwise. Definitions of all other variables are provided in the appendix. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.4: Summary Statistics for Breached and Control Firms**

| Panel A: Firm Characteristics | | | | | | |
|---|---|---|---|---|---|---|
| | Breached (A) | | Control (B) | | Difference (A-B) | |
| | (N = 122) | | (N = 122) | | (Mean) | (Median) |
| | Mean | Median | Mean | Median | $p$-value | $p$-value |
| Total Assets ($ million) | 15,501.24 | 4,128.90 | 11,103.88 | 3,289.49 | 0.24 | 0.47 |
| Tangibility | 0.21 | 0.12 | 0.22 | 0.13 | 0.81 | 0.53 |
| Profitability | 0.10 | 0.09 | 0.09 | 0.09 | 0.54 | 0.93 |
| Sales Growth | 0.10 | 0.07 | 0.06 | 0.06 | 0.10 | 0.69 |
| Leverage | 0.26 | 0.17 | 0.23 | 0.20 | 0.48 | 0.23 |
| Tobin's $q$ | 2.23 | 1.68 | 1.94 | 1.53 | 0.13 | 0.24 |
| CAPX/Assets | 0.04 | 0.03 | 0.04 | 0.03 | 0.72 | 0.85 |
| Cash Flow Volatility | 1.70 | 1.28 | 1.45 | 1.02 | 0.26 | 0.12 |
| Stock Volatility | 0.37 | 0.32 | 0.35 | 0.32 | 0.62 | 0.98 |
| Inst. Ownership | 0.75 | 0.76 | 0.74 | 0.76 | 0.88 | 0.87 |

| Panel B: Deal Characteristics | | | | | | |
|---|---|---|---|---|---|---|
| | Breached (A) | | Control (B) | | Difference (A-B) | |
| | (N = 289) | | (N = 284) | | (Mean) | (Median) |
| | Mean | Median | Mean | Median | $p$-value | $p$-value |
| Spread (bps) | 179.41 | 172.50 | 187.01 | 150.00 | 0.52 | 0.70 |
| Maturity (months) | 51.33 | 60.00 | 51.24 | 60.00 | 0.96 | 0.51 |
| Amount ($ million) | 1,005.56 | 600.00 | 904.13 | 450.00 | 0.47 | 0.01 |
| Covenants | 1.23 | 1.00 | 1.25 | 1.00 | 0.88 | 0.53 |
| Financial Covenant | 0.57 | 1.00 | 0.63 | 1.00 | 0.18 | 0.17 |
| Performance Pricing | 0.49 | 0.00 | 0.52 | 1.00 | 0.59 | 0.53 |
| Secured | 0.37 | 0.00 | 0.35 | 0.00 | 0.49 | 0.53 |

This table reports summary statistics for firms that experienced a data breach (*breached*) and *non-breached* (control) firms over the period 2005-2015. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, book leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. Panel A reports mean and median values for *breached* and control firm characteristics, as well as $p$-values for the test of difference in means and medians. Firm's characteristics are measured in the year prior to a data breach. Panel B reports mean and median values for loans of *breached* and control firms, and $p$-values for the test of difference in means and medians. Deal characteristics are measured in the four years prior to a data breach. Mean differences are measured using the $t$-test; median differences are tested using the Wilcoxon Mann-Whitney test.

**Table 1.5: Difference-in-Difference Loan Spread Regressions**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $\text{Breached}_{it}$ | 0.18*** | 0.17*** | 0.16*** | 0.16** | 0.16*** |
| | (0.07) | (0.07) | (0.06) | (0.06) | (0.06) |
| Maturity | 0.07 | 0.05 | 0.05 | 0.05 | 0.07 |
| | (0.06) | (0.07) | (0.06) | (0.06) | (0.07) |
| Amount | -0.22*** | -0.19*** | -0.19*** | -0.20*** | -0.20*** |
| | (0.04) | (0.03) | (0.04) | (0.04) | (0.03) |
| Secured | | 0.42*** | 0.38*** | 0.38*** | 0.36*** |
| | | (0.10) | (0.10) | (0.10) | (0.10) |
| Financial Covenant | | -0.11 | -0.10 | -0.10 | -0.09 |
| | | (0.07) | (0.06) | (0.06) | (0.06) |
| Performance Pricing | | -0.03 | -0.03 | -0.02 | -0.02 |
| | | (0.05) | (0.04) | (0.04) | (0.04) |
| SP Rating | | | -0.13*** | -0.13*** | -0.12*** |
| | | | (0.02) | (0.02) | (0.02) |
| No Rating | | | -1.55*** | -1.54*** | -1.37*** |
| | | | (0.26) | (0.26) | (0.29) |
| Credit Spread | | | | 0.08 | 0.04 |
| | | | | (0.09) | (0.10) |
| Term Spread | | | | -0.15 | -0.14 |
| | | | | (0.10) | (0.10) |
| Total Assets | | | | | 0.11 |
| | | | | | (0.09) |
| Tangibility | | | | | 0.38 |
| | | | | | (0.40) |
| Profitability | | | | | 0.10 |
| | | | | | (0.64) |
| Cash Flow Volatility | | | | | -0.04 |
| | | | | | (0.06) |
| Leverage | | | | | -0.33 |
| | | | | | (0.24) |
| Tobin's $q$ | | | | | -0.04 |
| | | | | | (0.05) |
| Stock Volatility | | | | | 0.45* |
| | | | | | (0.23) |
| Z-Score | | | | | -0.02 |
| | | | | | (0.03) |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.84 | 0.85 | 0.87 | 0.87 | 0.87 |
| Observations | 934 | 934 | 934 | 934 | 934 |

This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breaches on the cost of bank loans. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points. Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.6: Heterogeneous Treatment Effects by Type of Breached Information**

| | Financial | | Cyber | | Customer | | Repeated | | Records | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Breached$_{it}$ × Type | 0.33*** | 0.25** | 0.68*** | 0.60** | 0.24** | 0.14 | 0.04** | 0.02 | 0.17** | 0.12* |
| | (0.10) | (0.11) | (0.25) | (0.26) | (0.11) | (0.11) | (0.02) | (0.02) | (0.08) | (0.07) |
| Breached$_{it}$ | | 0.11* | | 0.15** | | 0.15** | | 0.12* | | 0.17*** |
| | | (0.07) | | (0.06) | | (0.07) | | (0.07) | | (0.06) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 |
| Observations | 934 | 934 | 934 | 934 | 934 | 934 | 934 | 934 | 934 | 934 |

This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breach types on the cost of bank loans. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points. Definitions of all other variables are provided in the appendix. All specifications include controls, loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.7: Difference-in-Difference Other Price and Non-Price Term Regressions**

| | Panel A: Other Price and Non-Price Terms | | | | |
|---|---|---|---|---|---|
| | Amount | Maturity | Lenders | Annual Fee | Upfront Fee |
| | (1) | (2) | (3) | (4) | (5) |
| Breached$_{it}$ | 0.06 | -0.02 | 0.08 | 0.10 | -0.36 |
| | (0.12) | (0.03) | (0.10) | (0.08) | (0.62) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.71 | 0.85 | 0.68 | 0.95 | 0.85 |
| Observations | 934 | 934 | 934 | 303 | 180 |
| | Panel B: Contractual Features and Covenants | | | | |
| | Cov. Intensity | Cov. Dummy | Perf. Pricing | Fin. Cov. | Secured |
| | (1) | (2) | (3) | (4) | (5) |
| Breached$_{it}$ | 0.24*** | 0.14** | -0.07 | 0.11* | 0.05 |
| | (0.07) | (0.07) | (0.08) | (0.06) | (0.05) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.73 | 0.65 | 0.57 | 0.64 | 0.75 |
| Observations | 934 | 934 | 934 | 934 | 934 |

This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breaches on other pricing and non-pricing loan terms. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variables are other pricing and non-pricing terms included in bank loan contracts. Definitions of all other variables are provided in the appendix. All specifications include controls, loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.8: Difference-in-Difference Covenant Type Regressions**

| | Asset Sweep | Debt Sweep | Equity Sweep | Dividend Restriction | Excess CF |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $\text{Breached}_{it}$ | 0.14*** | 0.13*** | 0.05 | 0.11* | 0.06* |
| | (0.04) | (0.05) | (0.03) | (0.06) | (0.03) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.72 | 0.71 | 0.70 | 0.73 | 0.68 |
| Observations | 934 | 934 | 934 | 934 | 934 |

This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breaches on bank loan covenant. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variables are covenant terms included in bank loan contracts. Definitions of all other variables are provided in the appendix. All specifications include controls, loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.9: Cross-Sectional Heterogeneous Effects by *Ex*-Ante Firm Characteristics**

| | Panel A: Profitability and Investments | | | | | |
|---|---|---|---|---|---|---|
| | Fortune 500 | | Profitability | | CAPEX | |
| | Yes | No | High | Low | High | Low |
| Breached$_{it}$ | 0.28*** | 0.08 | 0.16 | 0.21** | 0.25** | 0.04 |
| | (0.08) | (0.08) | (0.11) | (0.10) | (0.12) | (0.10) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.91 | 0.86 | 0.90 | 0.89 | 0.90 | 0.88 |
| Observations | 400 | 511 | 274 | 318 | 276 | 328 |
| | Panel B: Other | | | | | |
| | Cyber Risk | | Z-Score | | D/CE | |
| | Yes | No | High | Low | High | Low |
| Breached$_{it}$ | 0.14 | 0.36** | 0.32*** | 0.18* | 0.11 | -0.01 |
| | (0.08) | (0.13) | (0.12) | (0.09) | (0.12) | (0.08) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.89 | 0.92 | 0.90 | 0.91 | 0.90 | 0.88 |
| Observations | 583 | 186 | 226 | 339 | 315 | 377 |

This table reports cross-sectional OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breaches on the cost of bank loans. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The sample is restricted to firms with top and bottom tercile firm characteristics in the year before a data breach. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points. Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, firm, and year fixed effects. $p$-values of the difference of the interaction coefficient are reported. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.10: Cross-Sectional Heterogeneous Effects by Syndicate Characteristics**

| | Lead Share ($\alpha$) | | Syndicate HHI | | # of Lenders | |
|---|---|---|---|---|---|---|
| | High | Low | High | Low | High | Low |
| Breached$_{it}$ | 0.33** | -0.02 | 0.60** | 0.21 | 0.07 | 0.34*** |
| | (0.12) | (0.18) | (0.22) | (0.12) | (0.10) | (0.08) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.86 | 0.93 | 0.86 | 0.93 | 0.91 | 0.90 |
| Observations | 193 | 155 | 190 | 163 | 275 | 325 |

This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breaches on the cost of bank loans. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The sample is restricted to firms with top tercile and bottom tercile syndicate characteristics in the four years before a data breach. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points. Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, firm, and year fixed effects. $p$-values of the difference of the interaction coefficient are reported. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.11: Changes in Firm's Outcomes**

| | Panel A: Financial and Operating Leverage | | | | | |
|---|---|---|---|---|---|---|
| | Book Leverage | | Op. Leases | | Leverage + Leases | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\text{Breached}_{it}$ | 0.02* | 0.02* | 0.01** | 0.01** | 0.04*** | 0.04*** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Controls | No | Yes | No | Yes | No | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry-Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.92 | 0.93 | 0.97 | 0.98 | 0.94 | 0.94 |
| Observations | 1750 | 1750 | 1607 | 1607 | 1607 | 1607 |
| | Panel B: Profitability and Distress | | | | | |
| | ROA | | Z-Score | | Merton's DtD | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\text{Breached}_{it}$ | -0.02 | -0.02 | -0.85* | -0.78* | -1.02** | -0.87** |
| | (0.01) | (0.01) | (0.47) | (0.45) | (0.48) | (0.44) |
| Controls | No | Yes | No | Yes | No | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry-Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.63 | 0.65 | 0.85 | 0.86 | 0.79 | 0.81 |
| Observations | 1750 | 1750 | 1750 | 1750 | 1225 | 1225 |

This table reports OLS regression coefficients and standard errors from a difference-in-difference analysis for the effect of data breaches on firm outcomes. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The dependent variables in Panel A are book leverage ratio, operating leases, and total leverage. The dependent variables in Panel B are ROA, Altman Z-score, and Merton's distance to default. Definitions of all other variables are provided in the appendix. All specifications include firm, and industry-year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.12: Conditional Likelihood of Future Data Breaches**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| $\text{Breach}_t$ | 0.27* | 0.27** | 0.24* | 0.47*** | 0.44*** | 0.42*** |
| | (0.14) | (0.13) | (0.13) | (0.17) | (0.16) | (0.16) |
| $\text{Breach}_{t-1}$ | | 0.04 | 0.01 | | 0.16 | 0.14 |
| | | (0.13) | (0.13) | | (0.16) | (0.17) |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry Fixed Effects | No | No | Yes | No | No | Yes |
| Pseudo-$R^2$ | 0.03 | 0.03 | 0.04 | 0.07 | 0.07 | 0.08 |
| Observations | 1,840 | 1,840 | 1,840 | 606 | 606 | 606 |

This table reports coefficients and standard errors from probit regressions for the likelihood of a data breach given the firm has been breached before. The sample consists of firm-year observations from Compustat over the period 2005-2017. The dependent variable is an indicator variable that takes a value of one if a firm experiences a data breach in year $t+1$, zero otherwise. Columns (1) to (3) consist of all firms that suffered at least one data breach; columns (4) to (6) consist of all firms that suffered at least two data breaches. Definitions of all other variables are provided in the appendix. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.13: Loan Spread and Probability of Covenant Violations**

| | Prob. Violation | | Performance | | Capital | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\text{Breached}_{it}$ × Pr. Violation | 0.00* | -0.00 | -0.00 | 0.00 | 0.07** | 0.07** |
| | (0.00) | (0.00) | | | | |
| $\text{Breached}_{it}$ | | 0.13* | 0.14* | 0.12 | 0.10 | 0.10 |
| | | (0.07) | (0.07) | (0.07) | (0.07) | (0.07) |
| Pr. Violation | | 0.00*** | | | | |
| | | (0.00) | | | | |
| Performance | | | 0.00** | 0.00** | | 0.00** |
| | | | (0.00) | (0.00) | | (0.00) |
| Capital | | | | 0.01*** | 0.01*** | 0.01*** |
| | | | | (0.00) | (0.00) | (0.00) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.91 | 0.92 | 0.91 | 0.92 | 0.92 | 0.92 |
| Observations | 479 | 479 | 479 | 479 | 479 | 479 |

This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the heterogeneous effect of data breaches and covenant violations on the cost of bank loans. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points. The probability of covenant violation is computed using the methodology of Demerjian and Owens (2016). Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

### Table 1.14: Loan Mispricing By Lenders

| | Original | | | | Loan-by-Loan | | | |
|---|---|---|---|---|---|---|---|---|
| | Baseline | Abnormal | Placebo 1 | Placebo 2 | Baseline | Abnormal | Placebo 1 | Placebo 2 |
| Breached$_{it}$ | 0.16*** | 0.15** | 0.01 | 0.01 | 0.14** | 0.13* | -0.01 | 0.02 |
| | (0.06) | (0.06) | (0.13) | (0.2) | (0.07) | (0.07) | (0.07) | (0.07) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | No | No | No | No |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | No | No | No | No |
| Event-Pair Fixed Effects | No | No | No | No | Yes | Yes | Yes | Yes |

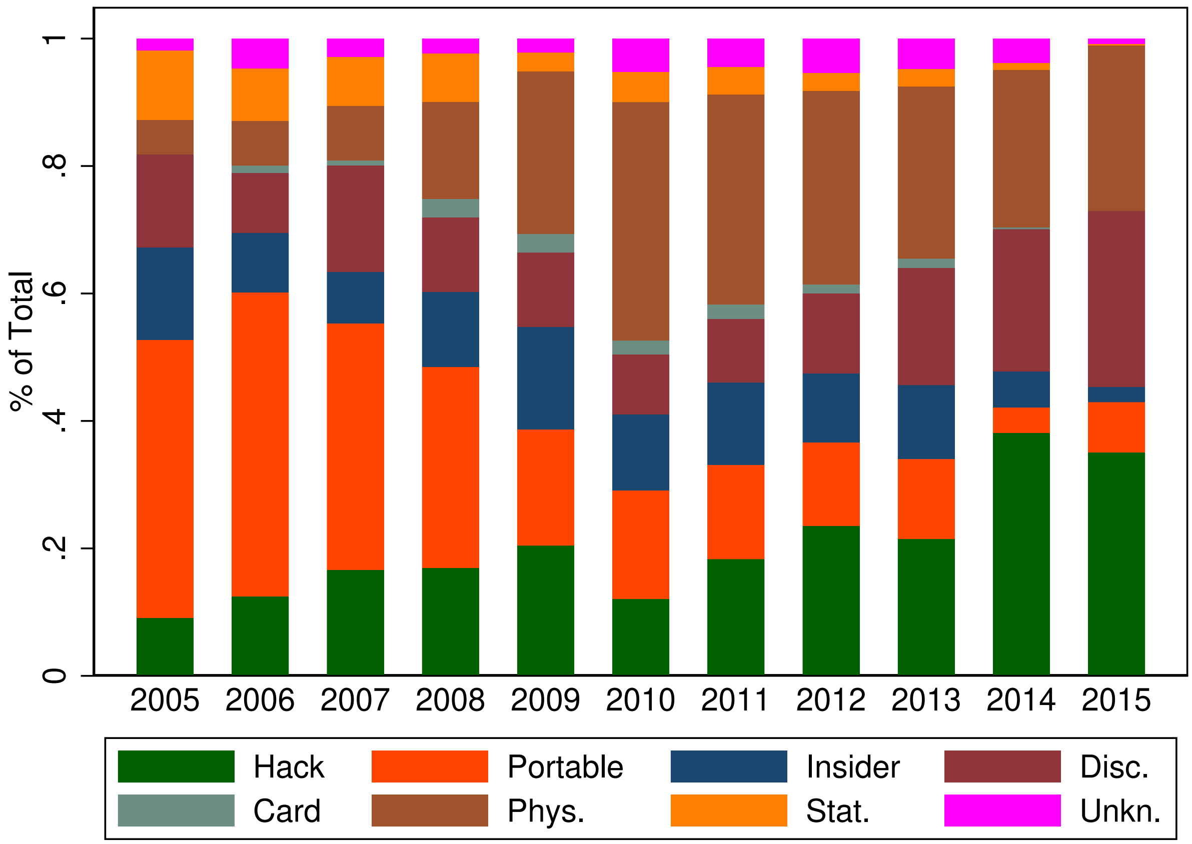
This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breaches on the *abnormal* cost of bank loans, and from two placebo tests. The sample of firms in the first four columns is the same as before, while the loan-by-loan sample consists of loans (treated) and matched loans (control) with synthetic event dates. Control loans are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, loan size, loan maturity, in the same 2-digits standard industrial classification (SIC) code, same year, same loan type and deal purpose. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variable is either the all-in-drawn loan spread over LIBOR (baseline), or the *abnormal* all-in-drawn loan spread over LIBOR in basis points, computed as the residual from a first stage regression of loan spreads on loan, firm characteristics, and fixed effects. Definitions of all other variables are provided in the appendix. Specifications may include loan type and purpose, firm, year, and event-pair fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

## 1.9 Appendix

**Effects on Corporate Bond Pricing and Non Pricing Features.** I documented in Section 1.3 that data breaches cause banks to charge higher spreads and modify some of the features included in loan arrangements. If banks can mitigate the effect of higher interest rates with covenants, institutional investors should charge similar (or larger) spreads. Larger spreads are expected, as there are no covenants or other stringent contractual terms in public corporate bonds. I obtain data on public debt from the Mergent FISD database to test the effect of data breaches on corporate bond pricing and non-pricing features. The main variable of interest is the at-issue-spread, calculated as the spread between the yield to maturity at issue and the yield to maturity of a government bond of similar duration. I linearly interpolate Treasury yields from the Federal Reserve Board to find matches for each public bond.

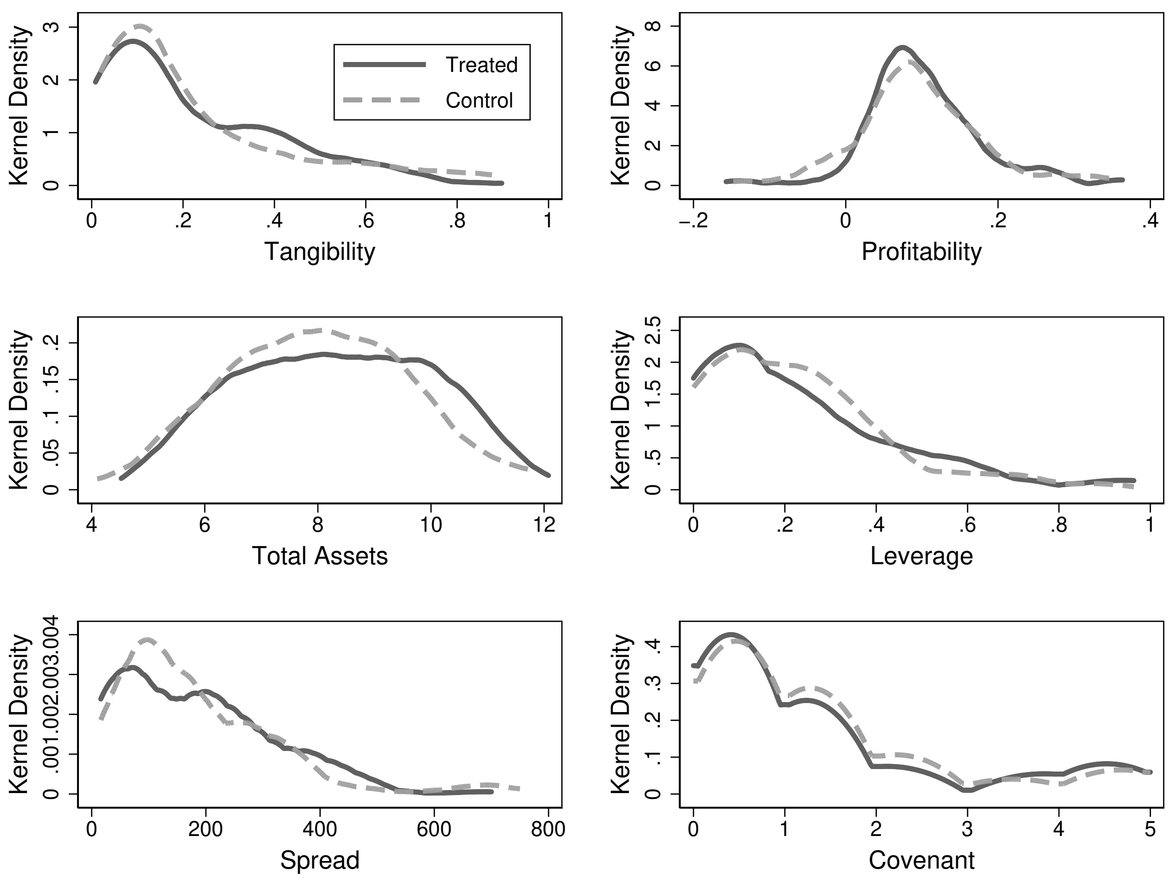
Table 1.20 shows results for the at-issue-spread over Treasury, maturity (in years), and total amount issued per bond. I control for year and firm fixed effects, as well as for the type of bond. Column (1) and (2) show that the effect on bond spread is large, although not statistically significant at conventional levels. However, its economic magnitude is almost double of that documented in the context of bank loans. Moreover, borrowers lengthen the maturity of their debt (consistent with Kamiya et al. (2018), and decrease the average amount per issues.

**Figure 1.3: Data Breaches by Type**




This figure shows the percentage of data breaches by type over the period 2005-2015 from the Privacy Rights Clearinghouse's Chronology of Data Breaches (PRC). PRC provides the following definitions of data breaches: CARD includes fraud (credit and debit card), not via hacking; DISC refers to unintended disclosure (not involving hacking, intentional breach or physical loss); HACK refers to hacking by a third party; INSD is someone with legitimate access intentionally breaches information - such as an employee, contractor or customer; PHYS includes paper documents that are lost, discarded or stolen (non electronic); PORT refers to lost, discarded or stolen laptop, PDA, smartphone, memory stick, CDs, hard drive, data tape, etc.; STAT refers to stationary computer loss (lost, inappropriately accessed, discarded or stolen computer or server not designed for mobility).

**Figure 1.4: Kernel Densities for Treatment and Control Firms**




These figures report kernel densities for treated and control firms the year before a data breach. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year.

**Figure 1.5: Pre-Trends in Bank Loan Outcomes**

This table reports average OLS regression coefficients from a deal-level difference-in-difference regression of the form:

$$Y_{i,t} = \gamma_i + \lambda_t + \sum_{\tau=-4}^{3} \delta_\tau Breached_{i,t+\tau} + \Lambda X_{i,t} + \epsilon_{i,t}$$

The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variables are pricing and non-pricing terms. Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. Confidence intervals represent a 90 percent range.

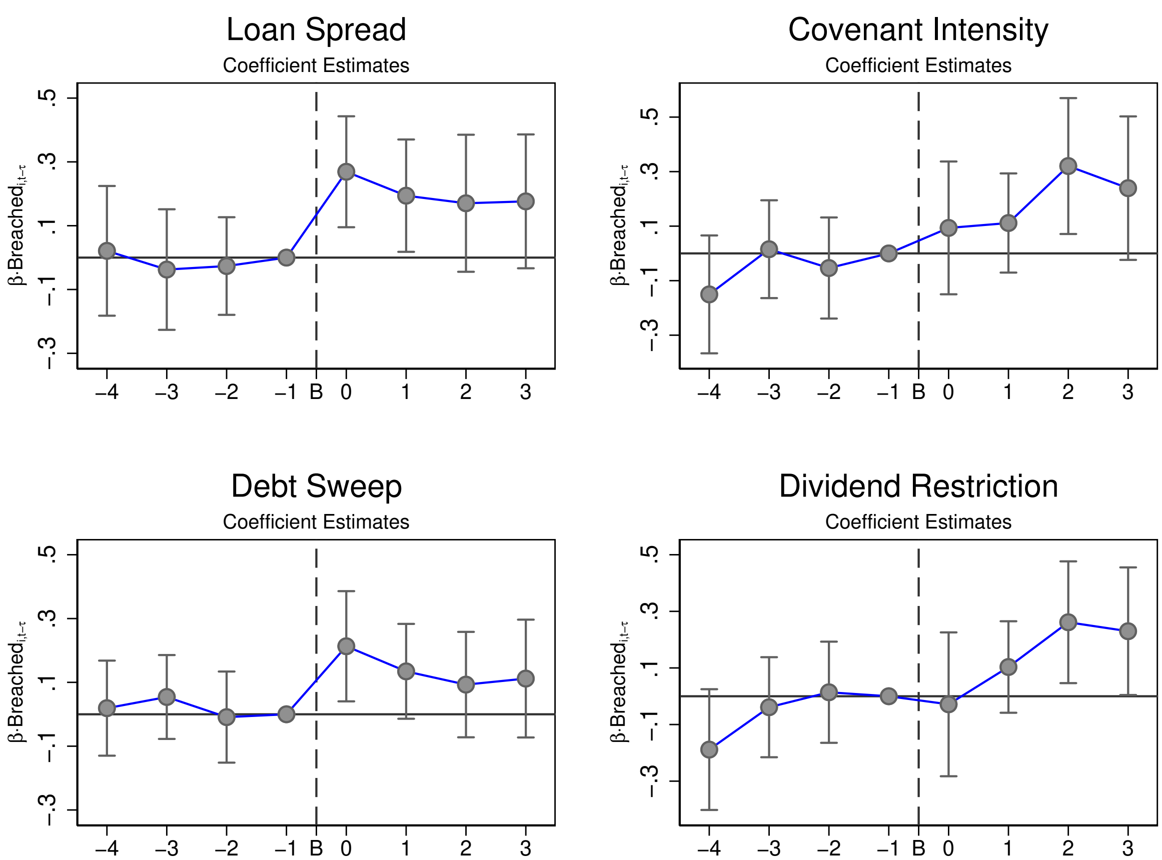


**Table 1.15: Single Difference Regression (as in Graham et al. (2008))**

| | Spread | Cov. Intensity | Fin. Cov. | Perf. Pricing | Secured | Amount | Maturity |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $\text{Post}_t$ | 0.20*** | 0.14** | 0.10 | -0.11* | 0.06 | 0.42*** | -0.00 |
| | (0.06) | (0.06) | (0.06) | (0.07) | (0.05) | (0.11) | (0.03) |
| Maturity | -0.07 | 0.03 | 0.07 | 0.11 | 0.08 | 0.78*** | |
| | (0.10) | (0.12) | (0.11) | (0.08) | (0.08) | (0.24) | |
| Amount | -0.19*** | -0.03 | 0.05** | 0.04 | -0.05* | | 0.06** |
| | (0.05) | (0.03) | (0.02) | (0.02) | (0.03) | | (0.03) |
| Total Assets | 0.17 | -0.20 | -0.18 | -0.01 | -0.04 | 0.02 | 0.18** |
| | (0.14) | (0.15) | (0.13) | (0.13) | (0.10) | (0.23) | (0.07) |
| Tangibility | -0.69* | -0.89 | -0.30 | -1.10 | -1.32** | -1.07 | -0.13 |
| | (0.38) | (0.76) | (0.83) | (0.84) | (0.61) | (1.10) | (0.36) |
| Profitability | -0.50 | -2.18 | 0.13 | 0.31 | -0.88 | -3.24* | 1.74** |
| | (1.22) | (1.39) | (1.24) | (1.17) | (0.97) | (1.90) | (0.86) |
| Cash Flow Volatility | -0.09 | -0.01 | -0.03 | 0.00 | 0.01 | 0.05 | -0.04 |
| | (0.07) | (0.10) | (0.08) | (0.10) | (0.06) | (0.11) | (0.04) |
| Leverage | 0.37 | -0.10 | -0.36 | -0.50 | 0.53 | -0.92* | -0.09 |
| | (0.29) | (0.41) | (0.34) | (0.34) | (0.37) | (0.52) | (0.18) |
| M-B | 0.00 | 0.00** | 0.00*** | 0.00*** | 0.00 | 0.00*** | -0.00 |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Z-Score | -0.04 | -0.01 | -0.03 | -0.03 | 0.04 | -0.03 | 0.01 |
| | (0.04) | (0.03) | (0.03) | (0.03) | (0.03) | (0.06) | (0.02) |
| Credit Spread | 0.05 | -0.01 | 0.06 | -0.06 | 0.00 | 0.11 | -0.05 |
| | (0.11) | (0.06) | (0.05) | (0.06) | (0.07) | (0.12) | (0.04) |
| Term Spread | 0.29*** | -0.02 | 0.02 | -0.00 | 0.03 | -0.12 | -0.05 |
| | (0.04) | (0.04) | (0.04) | (0.04) | (0.03) | (0.07) | (0.03) |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.86 | 0.70 | 0.59 | 0.52 | 0.74 | 0.73 | 0.88 |
| Observations | 471 | 471 | 471 | 471 | 471 | 471 | 471 |

This table reports OLS regression coefficients and standard errors from a deal-level single-difference analysis for the effect of data breaches on the cost of bank loans. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points. Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, and firm fixed effects, Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

Table 1.16: **Robustness Tests: 1 to 2 Matching**

| | Panel A: Loan Spread | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $\text{Breached}_{it}$ | 0.14** | 0.13** | 0.14** | 0.13** | 0.15*** |
| | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.85 | 0.86 | 0.87 | 0.87 | 0.87 |
| Observations | 1,275 | 1,275 | 1,275 | 1,275 | 1,275 |
| | Panel B: Contractual Features and Covenants | | | | |
| | Cov. Intensity | Fin. Cov. | Asset Sweep | Debt Sweep | Div. Restriction |
| | (1) | (2) | (3) | (4) | (5) |
| $\text{Breached}_{it}$ | 0.18** | 0.09 | 0.11** | 0.13** | 0.11 |
| | (0.07) | (0.06) | (0.05) | (0.05) | (0.07) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.72 | 0.63 | 0.67 | 0.67 | 0.61 |
| Observations | 1,275 | 1,275 | 1,275 | 1,275 | 1,275 |

This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breaches on the cost of bank loans and other pricing or non-pricing terms. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and two matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points or other pricing and non-pricing terms. Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.17: Robustness Tests: Different Matching Covariates**

| | Panel A: Loan Spread | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $\text{Breached}_{it}$ | 0.16*** | 0.15** | 0.15*** | 0.16*** | 0.17*** |
| | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.85 | 0.86 | 0.87 | 0.87 | 0.87 |
| Observations | 988 | 988 | 988 | 988 | 988 |

| | Panel B: Contractual Features and Covenants | | | | |
|---|---|---|---|---|---|
| | Cov. Intensity | Fin. Cov. | Asset Sweep | Debt Sweep | Div. Restriction |
| | (1) | (2) | (3) | (4) | (5) |
| $\text{Breached}_{it}$ | 0.17** | 0.13** | 0.10* | 0.14*** | 0.12 |
| | (0.08) | (0.06) | (0.05) | (0.05) | (0.07) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.70 | 0.60 | 0.65 | 0.63 | 0.59 |
| Observations | 988 | 988 | 988 | 988 | 988 |

This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breaches on the cost of bank loans and other pricing or non-pricing terms. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, profitability, and Tobin's $q$ in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points or other pricing and non-pricing terms. Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.18: Robustness Tests: Including Repeated Breaches**

| | Panel A: Loan Spread | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Breached$_{it}$ | 0.17*** | 0.18*** | 0.17*** | 0.17*** | 0.17*** |
| | (0.05) | (0.05) | (0.04) | (0.04) | (0.04) |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.85 | 0.87 | 0.88 | 0.88 | 0.88 |
| Observations | 1530 | 1530 | 1530 | 1530 | 1530 |
| | Panel B: Contractual Features and Covenants | | | | |
| | Cov. Intensity | Fin. Cov. | Asset Sweep | Debt Sweep | Div. Restriction |
| | (1) | (2) | (3) | (4) | (5) |
| Breached$_{it}$ | 0.16*** | 0.10** | 0.10*** | 0.12*** | 0.13** |
| | (0.05) | (0.05) | (0.04) | (0.03) | (0.05) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.72 | 0.63 | 0.65 | 0.64 | 0.63 |
| Observations | 1530 | 1530 | 1530 | 1530 | 1530 |

This table reports OLS regression coefficients and standard errors from a deal-level difference-in-difference analysis for the effect of data breaches on the cost of bank loans and other pricing or non-pricing terms. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of loans is from the DealScan database, originated in the four years before and four years after a data breach event. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points or other pricing and non-pricing terms. Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

## Table 1.19: Robustness Tests

| | Ownership | KZ-Index | WW-Index | Age | E-Index |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Breached$_{it}$ | 0.12$^{*}$ | 0.15$^{**}$ | 0.15$^{**}$ | 0.16$^{***}$ | 0.16$^{**}$ |
| | (0.07) | (0.06) | (0.06) | (0.06) | (0.07) |
| Inst. Ownership | -0.09 | | | | |
| | (0.33) | | | | |
| KZ-Index | | -0.00$^{*}$ | | | |
| | | (0.00) | | | |
| WW-Index | | | 0.02 | | |
| | | | (0.06) | | |
| Age | | | | -0.03 | |
| | | | | (0.19) | |
| E-Index | | | | | 0.04 |
| | | | | | (0.03) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Industry-Year Fixed Effects | No | No | No | No | No |
| $R^2$ | 0.88 | 0.87 | 0.87 | 0.87 | 0.88 |
| Observations | 786 | 901 | 908 | 934 | 748 |
| | Double Cluster | Time-Trend | Industry-Year | No Loan FE | No Year FE |
| | (1) | (2) | (3) | (4) | (5) |
| Breached$_{it}$ | 0.17$^{***}$ | 0.20$^{***}$ | 0.17$^{**}$ | 0.15$^{**}$ | 0.21$^{***}$ |
| | (0.06) | (0.06) | (0.07) | (0.07) | (0.06) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Loan Type & Purpose Fixed Effects | Yes | Yes | Yes | No | No |
| Year Fixed Effects | Yes | Yes | No | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Industry-Year Fixed Effects | No | No | Yes | No | No |
| $R^2$ | 0.87 | 0.87 | 0.92 | 0.84 | 0.84 |
| Observations | 934 | 934 | 853 | 940 | 934 |

This table reports OLS regression coefficients and standard errors from various robustness deal-level difference-in-difference tests for the effect of data breaches on the cost of bank loans. Institutional ownership is percentage of shares outstanding held by institutional investors, KZ is the Kaplan-Zingales Index, W-W is the Wu-Whited Index, Age is the natural logarithm of firm age in years, E-Index is the entrenchment index. The dependent variable is the natural logarithm of the all-in-drawn loan spread over LIBOR in basis points. Definitions of all other variables are provided in the appendix. All specifications include loan type and purpose, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.20: Public Corporate Bond**

| | Spread | | Maturity | | Amount | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Breached$_{it}$ | 0.36 | 0.31 | 0.20$^{*}$ | 0.16$^{*}$ | -0.28$^{**}$ | -0.18$^{*}$ |
| | (0.23) | (0.23) | (0.11) | (0.09) | (0.12) | (0.10) |
| Controls | No | Yes | No | Yes | No | Yes |
| Bond Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.73 | 0.74 | 0.34 | 0.35 | 0.48 | 0.54 |
| Observations | 756 | 756 | 756 | 756 | 756 | 756 |

This table reports OLS regression coefficients and standard errors from a difference-in-difference analysis for the effect of data breaches on the cost of public corporate debt. The sample of firms consists of firms (treated) that experienced a data breach over the period 2005-2015 and matched firms (control) with synthetic event dates. Control firms are matched using propensity score matching on the natural logarithm of firm total assets, leverage ratio, stock volatility, in the same 2-digits standard industrial classification (SIC) code, and same fiscal year. The sample of public corporate bonds is from the Mergent FISD database, originated in the four years before and four years after a data breach event. The dependent variable is the natural logarithm of the at-issue spread over treasury in basis points. Definitions of all other variables are provided in the appendix. All specifications include coupon type, security type, and Rule 144a, firm, and year fixed effects. Standard errors are adjusted for clustering at the firm-level. ***, **, * correspond to statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 1.21: Variable Definitions**

| Variable | Definition | Data Source |
|---|---|---|
| | Panel A: Firm Characteristics | |
| Total Assets | Calculated as the natural logarithm of a firm's total assets (at) | Compustat |
| Tangibility | Calculated as the ratio of property, plant and equipment (ppent) to total assets (at) | Compustat |
| Profitability | Calculated as the ratio of earnings before interest and taxes (ebit) to total assets (at) | Compustat |
| Return on Assets | Calculated as the ratio of net income (ni) to total assets (at) | Compustat |
| Sales Growth | Calculated as the yearly change in sales (revt) | Compustat |
| Book Leverage | Calculated as the ratio of short and long term debt (dlc+dltt) to total assets (at) | Compustat |
| Op. Leases | Calculated as the ratio of the present value of rental commitments using a 10% discount rate to total assets (at) | Compustat |
| Market to Book | Calculated as the ratio of the market value of equity (prcc_f x csho) to the book value of equity | Compustat |
| CAPX/Asset | Calculated as the ratio of capital expenditures (capx) to total assets (at) | Compustat |
| Age | Calculated as the natural logarithm of a firm's age in years | Compustat |
| Cash flow Volatility | Calculated as the annual standard deviation of firm's quarterly cash flows (ocfq) | Compustat |
| Altman Z-Score | Calculaetd as the ratio of $(3.3\times oiadp+0.999\times sale+1.4\times re+1.2\times wcap)$ to total assets (at) plus the ratio of $(0.6\times csho\times prcc\_f)$ to total liabilities (lt) | Compustat |
| Whited and Wu Index | Calculated as $-.091\times cf-.062\times divpos+.021\times tltd-.044\times lnta+.1021\times isg-.035\times sg$ | Compustat |
| Stock Volatility | Daily stock volatility over the fiscal year | CRSP |
| Distance to Default | Computed as $DD_t=[log(\frac{V_t}{D_t})+\mu-\frac{1}{2}\sigma^2]/\sigma$ where $V$ and $\sigma$ are estimated using Merton's structural model | CRSP/Compustat |
| | Panel B: Deal (Loan) Characteristics | |
| Loan Spread | Calculated as the natural logarithm of total annual spread over LIBOR for each dollar drawn down net of upfront fees | DealScan |
| Loan Amount | Calculated as the natural logarithm of the total loan amount in millions of dollar. The largest facility in a pacakge is retained | DealScan |
| Loan Maturity | Calculated as the natural logarithm of the loan maturity in months | DealScan |
| Secured | Dummy variable taking a value of one if the loan is secured by collateral, zero otherwise | DealScan |
| Financial Covenant | Dummy variable taking a value of one is the loan includes financial covenants, zero otherwise | DealScan |
| Performance Pricing | Dummy variable taking a value of one if the loan includes a performance pricing provision | DealScan |
| SP Rating | Standard and Poors rating dummies that take a value of 22 for the highest rating and a value of 1 for the lowest rating. A value of -1 is given if the rating is missing | Compustat |
| No Rating | Dummy variable taking a value of one if the firm is not rated, zero otherwise | Compustat |
| | Panel C: Breach Characteristics | |
| Breached | Dummy variable taking a value of one for breached firms in year 0, 1, 2, and 3, 0 otherwise | Privacy Rights Clearinghouse (PRC) |
| Financial Data | Dummy variable taking a value of one if the information breached contains financial data (credit/debit card, bank account, payment information) | Privacy Rights Clearinghouse (PRC) |
| Customer Data | Dummy variable taking a value of one if the information breached contains customer data (i.e any customer data of the breached firm) | Privacy Rights Clearinghouse (PRC) |
| Records | Total number of records divided by total assets (at) | Privacy Rights Clearinghouse (PRC) |

## REFERENCES


Akey, P., S. Lewellen, and I. Liskovich (2018). Hacking corporate reputations.

Almeida, H., M. Campello, B. Laranjeira, S. Weisbenner, et al. (2012). Corporate debt maturity and the real effects of the 2007 credit crisis. *Critical Finance Review 1*(1), 3–58.

Brown, G. W. and K. B. Toft (2002). How firms should hedge. *The review of financial studies 15*(4), 1283–1324.

Campbell, K., L. A. Gordon, M. P. Loeb, and L. Zhou (2003). The economic cost of publicly announced information security breaches: empirical evidence from the stock market. *Journal of Computer Security 11*(3), 431–448.

Campello, M. and J. Gao (2017). Customer concentration and loan contract terms. *Journal of Financial Economics 123*(1), 108–136.

Campello, M., C. Lin, Y. Ma, and H. Zou (2011). The real and financial implications of corporate hedging. *The journal of finance 66*(5), 1615–1647.

Carter, D. A., D. A. Rogers, and B. J. Simkins (2006). Does hedging affect firm value? evidence from the us airline industry. *Financial management 35*(1), 53–86.

Cavusoglu, H., B. Mishra, and S. Raghunathan (2004). The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. *International Journal of Electronic Commerce 9*(1), 70–104.

Chakraborty, I., A. J. Leone, M. Minutti-Meza, and M. Phillips (2018). Financial statement complexity and bank lending. In *29th Annual Conference on Financial Economics & Accounting*.

Chava, S., K. Huang, and S. A. Johnson (2017). The dynamics of borrower reputation following financial misreporting. *Management Science*.

Chava, S. and M. R. Roberts (2008). How does financing impact investment? the role of debt covenants. *The Journal of Finance 63*(5), 2085–2121.

Corielli, F., S. Gatti, and A. Steffanoni (2010). Risk shifting through nonfinancial contracts: effects on loan spreads and capital structure of project finance deals. *Journal of Money, Credit and Banking 42*(7), 1295–1320.

Demerjian, P. R. and E. L. Owens (2016). Measuring the probability of financial covenant violation in private debt contracts. *Journal of Accounting and Economics 61*(2-3), 433–447.

Deng, S., R. H. Willis, and L. Xu (2014). Shareholder litigation, reputational loss, and bank loan contracting. *Journal of Financial and Quantitative Analysis 49*(4), 1101–1132.

Gatzlaff, K. M. and K. A. McCullough (2010). The effect of data breaches on shareholder wealth. *Risk Management and Insurance Review 13*(1), 61–83.

Gordon, L. A., M. P. Loeb, and L. Zhou (2011). The impact of information security breaches: Has there been a downward shift in costs? *Journal of Computer Security 19*(1), 33–56.

Gormley, T. A. and D. A. Matsa (2011). Growing out of trouble? corporate responses to liability risk. *The Review of Financial Studies 24*(8), 2781–2821.

Graham, J. R., S. Li, and J. Qiu (2008). Corporate misreporting and bank loan contracting. *Journal of Financial Economics 89*(1), 44–61.

Ivashina, V. (2009). Asymmetric information effects on loan spreads. *Journal of financial Economics 92*(2), 300–319.

Jin, Y. and P. Jorion (2006). Firm value and hedging: Evidence from us oil and gas producers. *The Journal of Finance 61*(2), 893–919.

Kamiya, S., J.-K. Kang, J. Kim, A. Milidonis, and R. M. Stulz (2018). What is the impact of successful cyberattacks on target firms? Technical report, National Bureau of Economic Research.

Nini, G., D. C. Smith, and A. Sufi (2012). Creditor control rights, corporate governance, and firm value. *The Review of Financial Studies 25*(6), 1713–1761.

Rauh, J. D. and A. Sufi (2011). Explaining corporate capital structure: Product markets, leases, and asset similarity. *Review of Finance 16*(1), 115–155.

Smith, C. W. and R. M. Stulz (1985). The determinants of firms' hedging policies. *Journal of financial and quantitative analysis 20*(4), 391–405.

Spanos, G. and L. Angelis (2016). The impact of information security events to the stock market: A systematic literature review. *Computers & Security 58*, 216–229.

Sufi, A. (2007). Information asymmetry and financing arrangements: Evidence from syndicated loans. *The Journal of Finance 62*(2), 629–668.

Yuan, Q. and Y. Zhang (2015). Do banks price litigation risk in debt contracting? evidence from class action lawsuits. *Journal of Business Finance & Accounting 42*(9-10), 1310–1340.